What Makes the Shit Dope?
The Techniques and Analysis of Rap Flows

Kjell Andreas Oddekalv

Thesis submitted for the degree of Philosophiae Doctor
Department of Musicology
Faculty of Humanities
RITMO Centre for Interdisciplinary Studies in Rhythm, Time, and Motion
Shoutouts

It takes a village, a crew, a posse to create anything of substance—and if these 100000+ words contain something substantial, it is not because of me, but because a whole lot of people have pushed me along a long and winding path of thinking and typing these last years.

First of all, someone had to sign off on the brilliant idea of paying someone to nerd out about rap flows, and without RITMO, the Research Council of Norway, the Faculty of Humanities and the Department of Musicology at the University of Oslo, this thesis would never have seen the light of day. I am endlessly grateful for the trust these institutions have shown me.

At the centre of the Centre: The woman in charge—the gravitational centre of my academic universe, the single most important person for my scholarly endeavours, truly putting the ‘Super’ in front of ‘Star’—my supervisor Anne Danielsen. Without you, Anne, I would probably have been doing something completely different. From the very beginning of my master’s project and through the entire PhD you have been fundamentally inspirational. Not only have you been the best example any young scholar could fumblingly attempt to follow, but your pedagogical skills, razor sharp analytical mind and sincere enthusiasm for my project has been absolutely fundamental in shaping not only my work, but also me as a scholar and person. I am always immensely proud when telling people that you have been my supervisor, but I am even more proud of considering you my friend.

My co-supervisor Chris Stover is another hugely important contributor to this thesis. Your breadth of knowledge and the angles and viewpoints and attention to (minute!) details has been a real whetstone for sharpening my thinking and writing. Your similar-but-different approach to being a musician/scholar has been truly inspiring, and your philosophical and ethical considerations have also been invaluable to me. Not only for this project, but also for how I want to go forward with my academical work. Our discussions—including long digressions and the occasional sharing of craft beers—is something I have treasured and hope we will get to have more of in the future.

There are tons of others that I have been fortunate enough to cross paths with along this academic journey. Many have shown an interest in me and my work and shared of their valuable time without being obliged to. I want to mention Ragnhild Brøvig-Hanssen, my mentor at RITMO and yet another hugely inspiring scholar and lovely human being; Bruno Laeng, for reaching out with advise and feedback, completely unprompted; my colleagues in Forskningsgruppe for Sanglyrikk, particularly my rap-anthology co-editors Bjarne Markussen
and Even Igland Diesen; Venla Sykäri and Nigel Fabb from the Rhyme Conference and Sissel Furuseth from UiO for guiding me into the world of literary studies, literary linguistics and metrics; Mats Johansson, my midway assessment opponent and TIME-project colleague was crucial in making me embrace the ‘reflective practitioner’-angle; my copyeditor Nils; and there are the people in the European Hiphop Studies Network and many, many more. My colleagues at RITMO and the Department of Musicology (and in the TIME-project in particular) have been there along the journey, and their contributions are too many to be mentioned here. Suffice to say, I would never have gotten here without ALL of you, and I am immensely grateful for having been your colleague. I will mention only one in particular (although I could have written another thesis on all of these wonderful people), my ‘academic twin’ Bjørnar Sandvik: having you traversing your parallel, simultaneously very similar and very different path alongside me has been more valuable than you could imagine. Thanks for being there at conferences, in the office and online—a true partner in crime!

From ‘the real world’ my crews within the crew deserve some huge kudos as well. My homeboys in Sinsenfist—Bennærn, Ugstad, TeddyFoxXx, Reffrenget and Gutten—without you guys, I most definitely would not have been embarking on a journey to ratify my Dr.-status. The music and memories we create together is a bigger part of this thesis than is visible to most. Also, ‘Gjengen’ with a capital G—my very own G-unit—nobody could wish for a better supporting cast when writing your own life’s story (and a dissertation). And, of course, my family—mom, dad, brother, sisters, and the whole host of in-laws—you are the best kind of enablers, and I love you.

And finally, the most important ones, the nucleus, the atoms of my being. Iver—legendismens monument—you had not arrived when the journey began, but you made my world so much larger and any challenge so much easier when you showed up. Compared to your children, nothing else really matters, so writing a mere doctoral thesis became an insignificant task compared to being a father. Andreas, you were growing in your mama’s belly as I was finishing this thing—another brilliant motivator for getting shit done. I am proud of having written both this and other stuff, but that pride is completely insignificant in comparison to the pride I get from watching you learn to smile, laugh and—in due time—play with your big brother. Og Rachel… Jordingsledning, pærra-partner der vi begge to har fredning. På treet vi er, er eg greina og du stammen. Herrefred som eg elskar deg!

Kjell Andreas/Dr. Kjell/Kjellen, 2022
Tracklist

Shoutouts ........................................................................................................... ii
Tracklist ............................................................................................................. v
SIDE A

Intro ................................................................................................................. 2
Aims and research questions: ‘What makes the shit dope?’ .............................................. 2
Track A1: Flow, rhythm and analytical considerations .................................................. 7
Flow: What it can be, what it is, and what it is not .......................................................... 7
Rhythm: Flow’s direction and expression ...................................................................... 18
Context: The scholar, hip-hop and the world .................................................................. 22
Analysis: What we do, and how we do it ....................................................................... 30
Track A2: Rap analysis ......................................................................................... 34
The early days of rap analysis ...................................................................................... 34
The explosion of rap analysis ....................................................................................... 35
Corpus studies and a rap flow’s global features ........................................................... 36
Rhythmic surface structure: different schools of representation .................................. 38
Microrhythm, ‘expressive timing’ and (dis-)alignment between rhythmic layers .............. 42
Local form features: larger segments of flow ............................................................... 44
More than flow: the breadth of rap analysis .................................................................. 45
Four zoom levels of representation and analysis ........................................................... 48
Track A3: Rhythm cognition ............................................................................... 51
Streams, structures, schemata ..................................................................................... 51
Categorical rhythm perception .................................................................................... 52
Categorical rhythm perception and the zoom levels ....................................................... 55
Prediction, attention, markedness and expected convergence ..................................... 57
Summary ............................................................................................................... 60
Track A4: Metre on metre or verses in verses ......................................................... 61
Metre: Musical and poetic, experiential and analytical ................................................. 61
Metre on metre? Coinciding units or coinciding spans? ............................................... 64
The convention of line/bar coincidence: Should it dictate rap transcriptions? .......... 67
Lineation: What is a line? ......................................................................................... 73
Is rap ‘metrical’? ..................................................................................................... 76
Convergent and divergent metrical structure ............................................................. 78
Track A5: Rhyme and its rhythmic roles ................................................................. 82
Rhymes as boundary markers: Primary rhymes ......................................................... 86
Additional prominence from additional rhymes: Secondary rhymes ....................... 87
A short rhyme dictionary ......................................................................................... 88
Track A6: Levels of musical structure ................................................................. 95
Groove and song modes of listening .......................................................................... 95
Rap’s formal template: The verse .......................................................................... 99
SIDE B

Interlude ............................................................................................................. 131

Track B1: Divergent metrical structure and alternative lineations........................................ 134

Heteromorphies: Various interactions of primary rhyme, syntax and musical metre ............ 134
Bridge techniques .................................................................................................. 145
The long lines of André 3000 ................................................................................ 153
Summary................................................................................................................. 160

Track B2: Interacting with the gravity of the form .............................................................. 162

Grand gestures ........................................................................................................ 162
Cadencing ............................................................................................................... 166
‘Helt om natten, helt om dagen’: Lars Vaular’s mastery of gravity .................................... 170
‘F*ck Your Ethnicity’: working against symmetry .................................................... 175
Summary................................................................................................................. 180

Track B3: When timings meet: Microrhythmic techniques ................................................... 182

Metrical anchors: ‘Waterfalls’ .................................................................................. 182
Metrical anchors and rhythmic tolerance: ‘Si Eiga Rås’ ................................................ 186
Metrical anchors creating categorical ambiguity: ‘HiiIPower’ ......................................... 188
Bothness: ‘Eple’ ....................................................................................................... 192
Off and on-beat and stretched subdivisions: Bothnesses in ‘Do You Remember’ ............ 195
Do we really hear this stuff? Epistemological considerations ........................................ 198

Track B4: Speech-song or song-speech? ............................................................................ 202

Speech-song: Rhythmicity ......................................................................................... 202
99 problems, and is pitch one?..................................................................................... 204
Applying the ear test—is it rhythm, is it intonation, or is it Something Else? .................... 206
Utilising the in-between 1: ‘Casket Pretty’ .................................................................. 209
Utilising the in-between 2: ‘SONGS ABOUT YOU’ .................................................. 215

Track B5: What makes the shit dope? .............................................................................. 224

Convergences, ambiguities, dualities ............................................................................. 224
Research question 1: What are the core rhythmic features of rap flows? ....................... 228
Research question 2: How are rap flows analysed? ...................................................... 237
Rap analysis: What does it do, and what can it do? ...................................................... 240

List of figures ............................................................................................................ 244
Playlist .................................................................................................................... 247
Bibliography ............................................................................................................ 250
Throughout the thesis, I will scatter anecdotal musings which both contextualise and flavour the theoretical, epistemological and analytical discussions in the main text. They will be formatted like this paragraph—indented and italicised.

* 

For readability’s sake, references to recordings (parenthesised year of release) are only included the first time a track is referenced throughout. Since rappers very often go by monikers rather than birth names (and when birth names are used as stage names, the rappers’ first names are typically at least as prominent as their last names), the playlist (discography) is sorted alphabetically by first names and not last names (that is, ‘Kendrick Lamar’ is sorted under ‘K’, not as ‘Lamar, Kendrick’ under ‘L’).
Intro
Rap is everywhere. No matter where you go on planet Earth (or beyond, if there are humans there), you will probably hear *rap*. The musical practice of *rapping* has permeated the global membrane and cross-bred with local cultural expressions. Throughout the late twentieth and early twenty-first centuries, this form of expression has been adopted, appropriated and adapted over and over again. Its origins in African American *hip-hop* culture are often completely buried under layers and levels of recontextualisation, relocalisation and revitalisation. Sometimes, however, the impact of the pioneers and innovators of the late 1970s, 1980s and early 1990s is striking, as though time has stood still or the initial rap aesthetic has been treated like the drum break on a DJ’s turntable—repeatedly restarted and looped with each repetition shining some new light on otherwise familiar material. So, while rap is forever youthful, modern, hip and on the cutting edge, it is also aging, established, traditionalist and conservative. It is in many ways the ultimate cultural expression of the globalised world. It can accommodate new flavours or twists without sacrificing its central identity, and it can survive iterations which range from authentic, celebratory and culturally sensitive to commercialised, bastardised and even destructive.

What follows from rap’s cultural and artistic ubiquity is that writing about rap, at whatever level, is not a particularly original undertaking. This thesis, then, is more of a remix: yet another track reusing a sampled drum loop that others have sampled before. Still, while this thesis might seem to be treading familiar ground, part of rap’s magic is how well it communicates nuances, including those related to the personality, background and individual approach of the artist—or, in this case, the scholar. This thesis presents my take on a modern classic, the *analysis of rap flows*, and it will be different from any other because the breadth and depth of the subject outdo the breadth and depth of the musicological field. Through the lens of a particular Norwegian rapper/scholar (or scholar/rapper, perhaps), this thesis will look at the core questions involved in the rhythms of rap vocals as they unfold over time, in terms of the techniques and approaches rappers employ, and in particular how these techniques create and manipulate perceptual ambiguities. To do so, I must first unpack a lot of theory and terminology, so that is where we begin.

**Aims and research questions: ‘What makes the shit dope?’**

The goal of this thesis is twofold. First, it will provide insights into the aesthetics and structures of rap flows. Second, it will frame its analytical approach using a systematic
overview of the various existing approaches to analysing rap flows, including what those approaches offer and what they lack. The overarching research questions are as follows:

1: What are the core rhythmic features of rap flows?

2: How are rap flows analysed? That is, how have the answers to the previous question been produced, and how do these answers vary depending on the analytical framework in use?

Question 2 is simultaneously part of answering question 1 and a contextualisation of whichever answers I give to question 1, and it concerns both which tools have been used as well as what those tools’ underlying theoretical assumptions are. This dynamic also applies to the tools and theories introduced in this thesis. Addressing the questions in reverse order, the first part of the thesis contains a summary of the current state of academic rap-flow analysis, with a particular focus on the methods employed and the specific knowledge unearthed. As part of this review and discussion, I will try to systematise how these different tools and their application(s) shed light on certain core aspects of rap flows, outlining a sort of meta-theoretical framework for analysing rap flows and even the rhythms of other forms of musical expression.

This will set the stage for my own response to the first main question above, regarding the core features of rap flows. The sub-questions at this point are:

What characterises the structural framework(s) of a rap flow?

How do rap flows interact with and contribute to creating their inherent structural framework(s)?

What are the essential rhythmic techniques—structural and expressive—which rappers employ in this interaction?

These questions will be answered partly through a theoretical discussion based on existing theories and methodological frameworks of rhythm cognition and rap, music and poetry analysis, and partly through musical analysis. Here and especially in my own analytical approach, I will also draw upon my experiences as a performing hip-hop artist, as well as a ‘hip-hop scholar’, so that the hip-hop ethos and flavour informs the argument’s content, language and spirit.¹

¹ The term is taken from Rollefson (2017); see discussion in track A1 (page 32).
To explain how and why I ended up with these research questions, I must zoom out to the big initial question behind it all—in a way, the origin question for virtually all music analysis, and the title of the thesis as well. Fundamentally, music analysis is pointing out, describing and making sense of some sort of musical thing in an attempt to understand it better, learn how to reproduce it (or a variation of it) or express one’s appreciation of it. Invoking the discourse of hip-hop enthusiasts deconstructing their favourite tracks, we therefore ask:

\[ \text{Figure A0.1: Hand drawn mega-question with initial nuancing in blue pen.} \]

The research object of this thesis, the musical thing—‘the shit’ of the mega-question—is not rap as a whole, including its lyrical content, voice techniques, development of sub-genres and so on. Naturally, I am particularly fascinated by some of the shit (as implied by the parenthetical text in figure A0.1), and that is the rhythm in rap—rap’s flow. Thus, the mega-question narrows to what makes rap flows dope? To answer it, one needs to understand what a rap flow is, so another derivative question (shown by the large parentheses and the green colour of the word ‘dope’ in the figure) is what makes rap flows?, or rather what are rap flows made up of? It is also highly unlikely that I can cover everything compelling about rap flows in two hundred pages, so the prefix ‘some of’ is pertinent here too (again, as illustrated in the figure). A last question is what do I really mean by ‘dope’? This concerns which parts appear most significant to me when I make music myself, but it also invites discussion regarding certain aspects of what makes up a rap flow that are less researched and understood.

As should be clear from the table of contents, the structure of the thesis is organised like a vinyl record, with an A-side and a B-side comprised of a series of numbered tracks.

Side A is dedicated to introducing the epistemological and theoretical basis and frameworks of the thesis—both my own and those which I inherit from and share with other scholars. This side is mainly targeted towards answering the second main research question.
Side B is my presentation and analysis of the different techniques employed by rappers, and it is mainly targeted towards answering the first main research question and the sub-questions.

This first track (A1) is dedicated to basic terminological and epistemological issues. What is flow (and what is it not)? What is rhythm, and which aspects of it are in focus here? What is the place and role of this thesis, its topics and its author in the scholarly and musical world? What is analysis, and how should it be performed?

Track A2 is a review of the field of rap analysis, and particularly rap flow analysis. It is also an initial sorting of methodological tools into zoom levels to introduce a meta-theoretical framework for analysing rap flows. These zoom levels are, from high to low or wide to narrow, as follows: the global zoom level, the macrorhythmic zoom level, the quantised zoom level and the microrhythmic zoom level. Each is suited to the observation of a different set of rhythmic features.

In track A3, I present the concept of composite auditory streams and the theories of rhythm cognition on which it is based. I also present theories of categorical perception and how they relate to the zoom levels. In addition, I introduce the notion of expected convergence between elements of different layers of a compound auditory stream, framing it with theories of expectation/prediction.

Track A4 explores the relationship between musical and literary form and structure, and especially the interaction between musical metre and poetic lines. It is a theoretical exploration of the concepts of metre and lineation, as well as the theoretical and methodological implications of lyrical transcription. Here, my concepts of convergent and divergent metrical structure—two different types of relationship between musical bars and poetic lines—is introduced.

In track A5, I present the terminology and theory related to rhyme which is relevant within the context of rap flow analysis. This involves certain pragmatic binaries such as the choice to focus not on a typology of rhyme but rather on rhyme/non-rhyme, as well as the categorisation of rhyme classes’ structural importance as primary and secondary rhymes. This track also contains a rhyme dictionary of the terms encountered throughout the thesis.

In the last track on side A, track A6, I present and discuss theoretical concepts as they relate to different structural levels of rhythm—concepts involving different modes of listening, form, hypermetre and metre, stress and accents, and expressive timing and systematic
variation in microrhythm. I round off side A with a short analytical amuse-bouche showing how the different zoom levels can supply a framework for analysis, and how they interact when applied to a concrete musical example.

Turning over to side B, track B1 presents various rap techniques involving *divergent metrical structure* and *alternative lineations* and culminates with extended analyses of some of OutKast member André 3000’s flows showcasing these types of techniques.

Track B2 introduces the concepts of *grand gestures* and *cadencing*, showcasing how rappers work with or against the *gravity of the form* in the structuring of their flows. Extended examples of flows by Lars Vaular and Kendrick Lamar show contrasting approaches, with one actively emphasising hypermetric blocks within the larger verse form and the other actively working against hypermetric predictability.

Track B3 zooms in on the smallest of rhythms to present the concepts of *metrical anchors* and *bothness*. Each concept is related to the slight disalignment between categorical rhythmic information from different layers of the composite auditory streams, and my analyses show how rappers use this disalignment to create rhythmic ambiguity.

Track B4 discusses the boundaries between speech and song, and whether it is sensible to think of rap as a ‘between form’ of the two. The track is mainly focused on the rhythmic implications of *utilisations of the in-between*, but it also goes slightly outside of the thesis’s jurisdiction to explore certain melodic aspects of rap. It ends with two extended examples of rappers who utilise these speech/song boundaries in very different ways.

Finally, track B5 summarises the findings of the thesis and its responses to the research questions, then embarks on a discussion of the implications of its responses.
‘Flow’ is the central term in all rhythmic analysis of rap. It is simultaneously intuitively descriptive and surprisingly difficult to pin down, as its application is both broad and sometimes contradictory. Rappers, hip-hop ‘heads’2 and scholars all have their own definitions of the term, and few are in complete agreement. By engaging with a variety of scholarly definitions and descriptions, the next few pages will distil from them a definition which will be used throughout this thesis. In conjunction, I will discuss other dimensions and uses of the term to clarify my central accompanying terminology.

**Flow: What it can be, what it is, and what it is not**

Flow is a musical parameter, as Adam Bradley’s description of it as ‘the song a rapper’s speech sings’ (2009, p. 9) suggests. Slightly less poetic and more practical is his actual definition: ‘Flow is where poetry and music communicate in a common language of rhythm’ (Bradley, 2009, p. 5). Here, he emphasises a couple of key points—namely, that flow is rhythm, and that it is the rhythm found in the meeting of music and poetry. Tricia Rose’s description of flow introduces two other aspects, its *movement* and flow *as an ability*: ‘Rappers speak of flow explicitly in lyrics, referencing an ability to move easily and powerfully through complex lyrics, as well as of the flow in the music’ (Rose, 1994, p. 39). Alexs Pate also emphasises the forward movement inherent in the etymology of the word, but unlike Rose, he does not think of flow as an ability a rapper displays, but rather as something that is achieved through a rapper’s engagement with the structure of the rap:

> Flow is created from the effectiveness of the structure. Structure alone does not establish quality. But when flow is achieved, it is because the structure is present and establishes the environment in which the poet can flow. In other words, all of the images, rhymes and ideas held within the body of a rap/poem must coalesce into a whole. And the whole must move fluidly from beginning to end. This fluid movement—sometimes breakneck, sometimes languid—we call the flow. There are times when a rap/poet is deep within the rhyme pattern. The poem is alive with energy, and we know that ‘the poet’s flowin’’. (Pate, 2010, p. 115)

---

2 Hip-hop aficionados are often described as ‘hip-hop heads’ or simply ‘heads’.
Pate’s conception of flow recalls the term’s use in psychology—the idea of getting into a state of flow developed by Mihaly Csikszentmihalyi (1990)—which would imply that flow is a rap’s ideal state but not something which can be expected to appear in every rap.

One novel approach to defining flow which also emphasises the term’s etymological roots is that of Oliver Kautny (2015, pp. 103-104). He begins with a short definition—‘the rhythmic delivery of MCing is called flow’—before introducing ‘three dimensions of flow’: production (‘the air flowing out of the lungs, formed into a flow of sound’), texture (‘the musical result of the airflow synchronised to a musical arrangement called beat’) and reception (‘the feel of music while perceiving it’). The implication of the ‘rhythmic delivery’ from the short definition and the first dimension of ‘production’ seems to be that flow is something a rapper does, whereas the additional dimensions of ‘texture’ and ‘reception’ would indicate that flow is a feature of the music (that is, the result of ‘the rhythmic delivery of MCing’).

This dual nature is also present in other definitions. David Toop, Charise Cheney and Loren Kajikawa (2012), for example, write: ‘A rapper’s style of declamation, colloquially known as “flow”, imparts form and direction to the music by means of rhythmic patterns and rhyme schemes that interact dialectically with the instrumental track’. Once again, flow is something tied to the rapper (their declamatory style), but its constituent parts are features of the musical output (rhythmic patterns and rhyme schemes). Like Kautny, Toop, Cheney and Kajikawa stress the relationship between the rapper’s voice and the musical background (they use ‘instrumental track’ whereas Kautny uses ‘musical arrangement’ and the colloquial term ‘beat’). This is also reminiscent of Pate’s focus on the structure surrounding the flow. Flow, then, always interacts with something.

This something is not necessarily the sounding music, however. While rap is archetypically accompanied by a musical background—an instrumental ‘beat’—it is also routinely performed a cappella, or without musical accompaniment. The latter iteration does not imply a lack of musical structure, however, as is evident in Mitchell Ohriner’s structured definition of flow (2019b, Chapter 1), which he derives from the ways rappers ‘flow about flow’—or, more precisely, from the answers to the question ‘what do emcees do musically when they are flowing?’ (2019b, p. 16). Ohriner identifies many aspects of flow through his analysis of MCs

---

3 That a musical feature is both produced and perceived is true of most musical parameters in most forms of music-making, so this is not something which is unique to flow. However, it is a good argument for flow being a musical parameter rather than exclusively the product of a rapper’s ability, for example.

4 Note that in colloquial hip-hop discourse, ‘a cappella’ tends to refer to isolated tracks of rap vocal recordings, whether or not they were originally performed, recorded or intended as such.
referencing a ‘flip of the flow’ in their lyrics, then refines them into the ‘primary and derived constituents of flow’, where the former—words, phrasing and rhythm—create the latter—rhyme, accent and groove. From all of this, he defines flow as ‘the interaction between the derived constituents and meter’. I take some issue with the hierarchical ordering but especially the term ‘derived’, as it implies a power relation in these constituents’ interaction, whereas I propose that Ohriner’s ‘derived constituents’ create the primary constituents as well. Additionally, even if a rapper were capable of isolating individual constituents, the choice of which ‘primary constituents’ to use in a flow would clearly depend upon one or more of its ‘derived constituents’ in turn. It is worth noting that Ohriner’s concept of ‘groove’ (or ‘vocal groove’) is also unique to his work and quite different from most other definitions or understandings of the term. He uses ‘vocal groove’ to describe a specific type of accent patterning within cycles of sixteen sixteenth notes (that is, one bar of music). Different ‘vocal groove classes’ are made up of different sets of inter-accent-intervals, expressed by numbers like ‘[233233]’ (Ohriner, 2019b, p. 83). While this definition must be unpacked in the context of his work, which I lack the space to do here, I welcome his emphasis upon rhyme and accent as important parts of flow. I also appreciate the fact that Ohriner identifies metre as the musical framework with which flow interacts—other definitions point more generally to the musical background/instrumental track instead (metre as a concept will be extensively discussed in tracks A4 and A6). Since his definition emphasises the interaction between rhyme (as one of the three ‘derived constituents’) and metre, one could say that his definition of flow encompasses rhyme position.

At this point, what seems to be agreed upon is that flow is fundamentally rhythmic, and that it is communicated through and emerges from the rhythm of the words in the rap and how those words interact with some sort of musical rhythmic structure. The divisive point is whether flow is a feature of the rapper (a stylistic identity) or a feature of the rap (a rhythmic structure). In the latter camp are definitions such as Martin Connor’s—‘[Flow] means the rhythmic structure that arises in a rap from the interaction between the rapper’s words and the strictly musical rhythms of those words’ (2013)—and Paul Edwards’s—‘The flow of a hip-hop song is simply the rhythms and rhymes it contains’ (2009, p. 61). Conversely, Adam Krims defines flows as ‘rhythmic styles’ which can be tied to both specific rappers and larger groups:

The rhythmic styles of MCing, or ‘flows’, are among the central aspects of rap production and reception (…) Rhythmic style marks several dimensions of rap music
at once for artists and fans—history, geography, and genre all at once, not to mention
the constant personal and commercial quest for uniqueness. ( . . . ) [T]here is little
common language that I have been able to observe to describe flow. Just about any rap
artist or fan that I have encountered has well-developed notions of (and preferences
for) rhythmic style; but artists and fans tend to refer those notions sometimes to
particular artists, sometimes to history or geography, even sometimes to particular
songs. So many different frames of reference have rendered popular terms for flow
diverse to the point where generalization becomes confusing. (Krims, 2000, p. 48)

While more than twenty years have passed since Krims described ‘flow’ as frustratingly
multifaceted, there is still no real consensus around the term. Both colloquially and in
academic discourse, it is used (primarily) in two different ways. Either it refers to musical
rhythmic structure or it refers to a ‘style’. 5 While this can be frustrating in scientific contexts
where precise terminology is paramount, it can also be a boon if one is curious about the
depth and nuance of a complex concept. Likewise, the ambiguous usage is seemingly
unproblematic for fan discourse or the popular media, where the context informs which
meaning of the term is in play, and the difference rarely matters much in any case. My two-
part solution for determining a working definition of ‘flow’ is (1) to choose the definition
which is most frequently implied in the contexts I occupy as a rapper (the studio environment,
for example), and (2) to read some uses of the term as having silent suffixes, whereby ‘flow’
might be used as a shorthand for ‘flow style’ or ‘flow type’.

Beginning with the latter point, formalising Krims’s notion of flow as ‘flow style’ seems to
make sense, as it is both a part of the definition ‘rhythmic styles of MCing’ and an indication
that there can be a taxonomy of levels of flow. For example, it could refer to the personal
style of a particular rapper, the style of a group consisting of several rappers, the style of
rappers from a certain geographical area, or the sort of larger general category which Krims
himself presents. In short, Krims’s taxonomy of categories divides all rap into three flow
styles: ‘sung’, ‘speech-effusive’ and ‘percussion-effusive’ (Krims, 2000, pp. 49-51). These
styles are differentiated by their rhythmic organisation (that is, the amount of syncopation or
off-beat phrasing, for example) and the relationship between musical metre and poetic

---

5 Krims hints at this duality in the quote above when describing how rhythmic style can be found even in
particular songs.
structure within them, as well as by the way the emcees\textsuperscript{6} use their voices (as the names of the categories indicate). While I will return to Krims’s taxonomy of flow styles in track B4 and occasionally refer to rappers’ flow styles in my analyses, the important point for now is that I consider ‘flow style’ to be something other than ‘flow’.

I also want to differentiate flow with another silent suffix—‘type’—from ‘flow style’. An important aspect of the concept of ‘flow styles’ is that they communicate something about the identity of an individual or group. While flow types, like styles, are categories defined by certain parameters, they are unlike styles in that they are defined by specific features of the musical rhythm and can feasibly be employed by rappers with differing stylistic identities. Examples of flow types include ‘triplet flow’, a term commonly used to refer to flows popularised by artists like Migos\textsuperscript{7} wherein the rap is mostly (and often exclusively) subdivided into triplets,\textsuperscript{8} and ‘stutter rap’, which is characterised by short groups of syllables (often two, but occasionally more) placed on the beat with clear pauses on the off-beats.\textsuperscript{9} While these flow types are often tied to more holistic stylistic identities, they do not need to be, and they are also employed in contexts (and by rappers) where they might be unexpected or unfamiliar.

My final argument for splitting the flow term in three—flow, flow [style] and flow [type]—is anecdotal. When discussing, writing and recording rap, my peers and I will tend to use ‘flow’ to describe rhythmic structure: ‘I have the lyrics, but I’m figuring out the flow at the end there’; ‘what if you changed the flow to get that syllable on the final beat?’; ‘perhaps the flow would be less frantic if you start that line before the one’. When listening to a track and exclaiming, ‘damn, what is going on with the flow there’ or ‘that was some dope flow’, our common understanding of the notion is that it refers to something very specific, measurable and repeatable which is part of the musical structure.

Ultimately, ‘flow’ will carry with it all of these nuances and contradictory multiple meanings whichever way one chooses to define it, but for the context of this thesis, I will use the following:

\textsuperscript{6} There are different conventions regarding how to represent the abbreviation for ‘master of ceremonies’ in writing. I prefer using the spelled-out ‘emcee’ to emphasise the fact that the term has taken on many different layers of meaning since its inception.
\textsuperscript{7} In fact, ‘Migos flow’ is often used synonymously with triplet flow, because the Atlanta outfit used it so prominently andiconically when they first arrived on the scene.
\textsuperscript{8} See Duinker (2019) for a discussion of ‘modern’ triplet flow. Triplet flow is also a popular topic for YouTube videos and other popular scientific public-media channels.
\textsuperscript{9} See Komaniecki (2019, pp. 33-35) for a discussion of stutter rap with examples.
Flow is the rhythmic structure of the words and rhymes in the vocal track(s)/performance in a work of rap music.

With this definition, the primary aspects of flow are represented and the potential openness of some of these aspects is also preserved. The formulation ‘vocal track(s)/performance’ is intended to acknowledge that a rap flow *can* be performed live, but that most analysis of and other interaction with rap flows involves recorded music. And in a recorded context, the rap flow is rarely an unedited ‘raw’ performance. Instead, it is typically altered—it is cut together from different *takes*, for example, and/or several tracks have been stacked to emphasise certain passages or single words in so-called backtracks or dub tracks. While the edited and mixed recording can (and probably should) be considered a ‘performance’ in a wider sense of the word, it remains true that a rap flow on a recording is (most of the time) not just a single voice captured and preserved via microphone.

The reasoning behind including ‘in a work of rap music’ in the definition rather than leaving it at ‘in rap music’ is to emphasise the importance of the relationship between the flow and the formal unit(s) of which it is a part. While the flow itself is a significant aspect of creating the musical form, that formal unit is also communicated/created by the other parts of the musical context, and it has a significant impact on the rhythmic structure of the flow. The most common formal unit of a rap flow is the *verse*, or what the poets and literary scholars would refer to as a *stanza*, but it can also be any other structural unit with some sort of beginning and end. There are rapped choruses/‘hooks’, intros and outros, in addition to the archetypical verses. It is also possible to subdivide formal units, and these excerpts have flows even when they have been taken out of their complete original context. The trope of ‘flipping the flow’, for example, can be thought of as different sections of a verse having ‘different flows’.

Similarly, many analyses of rap flows will naturally focus on small segments of larger formal units, and these segments will still be ‘flows’. This means that by including ‘a work of’, the definition retains an emphasis on flow’s dependence on its interaction with a structural framework.

Some inclusions and omissions of constituents in my working definition require justification. Not all definitions of flow include rhyme, for example, but mine, like Ohriner’s and Toop,

---
10 This stance of treating the recording as a ‘performance’ is taken by, among others, Rollefson, who argues that ‘we can embolden our disciplinary move past the musicological ideology of the “musical object” toward an understanding of music as performance—even if that performance is crafted, (temporarily) fixed, and etched in a musical score or the grooves of a vinyl LP, or digitally encoded into an mp3’ (Rollefson, 2017, p. 10).
Cheney and Kajikawa’s, does. While most rap-flow analysts analyse rhyme in some way, particularly how rhymes are positioned within the form, there is an argument to be made for rhyme being outside of flow. Jakob Schweppenhäuser (2014, pp. 141-142) insists that rhymes should be considered ‘a different layer of rhythm’ than the rhythm of the syllables that constitute flow; if one were to remove the rhyme from a rap passage, he points out, the flow would remain the same. While this position is pragmatically reasonable (for some types of analysis), I disagree with it. The rhythmic role of rhyme in rap is so significant that I believe it should be included in the definition of flow, both when it comes to segmentation into phrasal units (lines, which will be thoroughly discussed in track A4) and when it comes to creating prominence—that is, making rhyming words/syllables stick out from their surroundings (a topic for track A6). Rappers organise their flows around rhymes, in short, making rhyme critical to how a flow is both experienced by the listener and assessed by the performer or expert listener. In fact, rhyme’s role in rap is very clearly more than its contribution to flow (something which will be discussed further in track A5), so a crucial point in my definition is that it is the rhythm of the rhymes which is part of flow.

One way my definition differs from some others is that I omit any mention of melodic and timbral features. Kyle Adams, for instance, writes that “flow” describes all of the rhythmical and articulative features of a rapper’s delivery of the lyrics’ (2009), then presents two categories of techniques: ‘metrical’ and ‘articulative’. The former are features of the rhythm of the words and rhymes and thus within the bounds of my definition: the position of rhymed and accented syllables, the relationship (‘degree of correspondence’) between syntactical and musical metrical units, and rhythmic density (‘the number of syllables per beat’). The latter, on the other hand, are more difficult to situate within my definition: the ‘amount of staccato and legato used’, the ‘degree of articulation of consonants’ and the ‘extent to which the onset of any syllable is earlier or later than the beat’. All of these can be considered ‘rhythmic’ in the sense that they influence our rhythmic experience in some way, but they do not necessarily contribute to a rhythmic structure. Instead, the rhythmic role of these articulative features tends to be expressive rather than structural, and while this is a distinction of great nuance (and this nuance is a considerable part of this thesis), I consider rhythmic expression to be part of another constituent of rap: namely, delivery.

The distinction between flow and delivery, and a third constituent of rap—content—is well presented in Paul Edwards’s seminal How to Rap: The Art and Science of the Hip-Hop MC (2009), a book based on interviews with many American rappers aiming to teach its readers
about the processes, mindsets and techniques employed by professionals. It is divided into sections based on these three taxonomical constituents (plus one on writing). The categories are presented by Edwards in the beginning of chapter 1 (about content):

The content of a hip-hop song (sometimes called the subject matter) includes every subject you talk about in your lyrics. It is what you’re actually rapping about, rather than the rhythms and rhymes you’re using (the flow), or how you’re using your voice to perform, or ‘spit’, those rhythms and rhymes (the delivery). (Edwards, 2009, p. 3)

Note that Edwards (and the rappers he interviews) uses the term ‘delivery’ differently from some of the scholars presented earlier. When Kautny writes ‘the rhythmic delivery of MCing’, for instance, he refers to the rhythm of the syllables in a rap performance, or what Edwards would consider flow. The use of the word ‘delivery’ in this instance is mainly pointing out that the flow is produced by a rapper; it does not engage with Edwards’s conception of the term. Adams’s ‘articulative techniques’, on the other hand, overlaps extensively with some of Edwards’s constituents of delivery. Edwards also mentions pitch, both in a stylistic-expression sense of being generally high- or low-pitched or varying a lot or a little and in the melodic sense of matching musical pitches with the instrumental track and/or singing shorter or longer segments. Some scholars explicitly include pitch in their definitions of flow. Gilbers, Hoeksema, de Bot, and Lowie (2019) define flow as the ‘rhythmic and melodic aspects of a rap performance’, and the previously mentioned Schweppenhäuser (2014) also argues for pitch’s inclusion as a constituent of flow. I consider rap’s melodic parameters to be (perhaps the most significant) part of delivery and thus mostly separate from flow, and I will again turn to practical anecdotes as my justification.

The two most intuitive contexts for rap being delivered are in the recording environment and when performing live, and both frequently feature situations where one very consciously keeps the flow the same while altering the delivery. A common point of discussion when someone is recording a verse is the melodic delivery. The rapper on the mic might ask ‘should I do it “up” or “down”?’, referring to a choice between two octaves. Similarly, the rapper might try a couple of different melodic variations on a section or try recording the verse with a more relaxed or more aggressive tone. All these different versions and different deliveries have the same flow. Along these lines, varying the delivery while keeping the flow the same

---

11 At least, this is my interpretation of Kautny. His second aspect of flow—‘texture’ encompassing ‘the musical result’ of the performance—might also include melody and some features of articulation which I would consider part of delivery, however.
is something my bandmates and I often do when performing live. After rapping the same
verse for years, I can ‘keep it fresh’ by altering one aspect slightly—for example, singing
parts of it (with slightly different melodies at different concerts) or doing my best Snoop
Dogg impersonation and rapping as laidback and relaxed as possible. Generally, I would
argue that, in terms of its production—perception is another matter—a rap’s flow is
something that is ‘written’ in the sense of composed. While many rappers will go on the mic
with some parts of their flow ‘to be determined’, they will typically have a general idea of
what they are going to do. Even fully ‘freestyled’\(^\text{12}\) rap recordings are ‘written’ in the sense
that their repetition, in live performance or via the playing of a recording, is a formalisation or
encoding of their rhythmic structure.

The examples of recorded freestyled rap and a ‘Snoop Dogg–like delivery’ touch upon a point
which will recur throughout this thesis—that is, the floating boundaries between rhythmic
structure and rhythmic expression, and consequently between flow and delivery. While there
are rhythmic features which might be considered expressive rather than structural in isolation,
they can become structural if they are systematic enough. Through repetition, that which first
appeared to be expressive delivery becomes inscribed into the structural flow. Other
contextual clues might also make the expressive structural. The listener’s expertise, focused
attention or familiarity with the artist’s style might inform the listening. The topic of the
systematically expressive becoming structural, particularly as it pertains to microrhythm
(identified by Adams as ‘articulative’), will be thoroughly revisited in track A6.

While this taxonomy is useful in clearly distinguishing among flow, content and delivery, it is
equally important to recognise and explore the fact that these categories are interdependent
and interactive. Content and delivery are not part of flow, but they can be strong contributors
to it. This nuance is important, and to unpack it, I must reentangle the disentangled categories
and discuss how they influence one another.

Beginning with content, it is fairly easy to understand why it is considered to be distinct from
flow and delivery. Flow and delivery are what the rap sounds like, while content is what it
means, and the former are intuitively separable from the latter. To anyone who has listened to
even a little rap music, it is obvious that the rhythms of the vocal track or performance can be

\(^{12}\) The term ‘freestyle’ is typically used for improvised rap, even if the term originally referred to rap that did not
aesthetically pleasing even when one does not pay attention to the actual lyrical content. When PSY’s ‘Gangnam Style’ conquered the world in 2012, one did not need to understand Korean to enjoy his rapping. There are also examples of entire subgenres or musical movements where the content of what is rapped is relatively unimportant to the music’s aesthetic—take, for example, the extensive use of not exactly poetically ambitious rap in the ‘Eurodance’ music of the 1990s and 2000s. Another example of content being set aside in the interests of prioritising delivery and flow is the so-called mumble rap popularised by artists like Gucci Mane, Chief Keef and Future. Lyrical incoherence, or at least a very indistinct enunciation, is perhaps the defining feature of the entire subgenre. And while this style might be slightly controversial to some older heads, other rappers who might themselves be considered quite lyrically ‘deep’ and proficient are enthusiastic about it, including the Norwegian rapper Linni:

[Gucci Mane] has always been inspiring. When he started to be inarticulate and made it sound better than being articulate, it was, like, ‘what a damn insight’! To, like, get that vibe comes first, and prove that you can say, ‘Peter Piper picked a peck of pickled peppers’ twenty times in a row, in a way. (Linni, 17.01.19)

What mumble rappers understand is that the semantic content of words and what the words sound like are two different things. Phonemes combined into morae and syllables are as much musical sounds as they are conveyers of semantic content. You need rhymes to rap, and you need syllables to rhyme, but the flow does not need to be impacted in any way by whether those syllables convey a political statement or a profane pun or nothing at all. Even within the discipline of written poetry, the separation of sound and content is a significant concept, and (as can often be the case in rap) the composition of the rhythm might well precede the words. This is beautifully explained by Northrop Frye (1957) in his exploration of the Aristotelian term melos and his translation of it as babble:

---

13 Several scholars have also discussed situations where the appreciation of the music is despite the lyrical content rather than because of it. Both Rose (1994) and Krim (2000) problematise the lyrical content of a significant part of the rap corpus they describe. Similarly, a decade later, Pate (2010, in particular chapter 4) does a great job of discussing some of the more nefarious aspects of certain rap lyrics (such as profanity, the ‘n-word’, misogyny and homophobia).

14 Linni is the moniker of Bergen native Jonas Grieg, known from the legendary rap collective Yoguttene. As a solo artist and in the duo project Neste Planet (alongside producer Kvam), he has been among the most prolific recording rap artists in Norway from around 2017 to the time of writing. The interview was conducted by the author, transcribed by Eirik Jacobsen and translated by the author.

15 Some rappers, in fact, write their flows before their lyrics (Edwards, 2009, pp. 113-114).
In babble, rhyme, assonance, alliteration and puns develop out of sound-associations. The thing that gives shape to the associating is what we have been calling the rhythmical initiative, though in a free verse poem it would be rather a sense of the oscillations of rhythm within an area which gradually becomes defined as the containing form. We can see from the revisions poets make that the rhythm is usually prior, either in inspiration or in importance or both, to the selection of words to fill it up. (Frye, 1957, p. 275)

But while a study of rap flows might in its very conception mean treating lyrics as *melos*/babble—as flow and delivery separated from semantic content—there are some exceptions to this rule of which we must be mindful. First, lyrical content often goes hand in hand with aesthetic choices regarding how flow is structured, in the sense of ‘party content is a good way to show off an expressive flow’ (Edwards, 2009, p. 21). Dense or serious lyrical content might get interfere with the rhythmic expression, or the other way around. Perhaps more pertinent is semantic content’s influence on prominence (how much a word or syllable ‘sticks out’ from its surroundings) and linguistic segmentation. While semantic content might be less important than either syntax (word ordering/sentence structure) or phonology (the patterning of speech sounds) for these parameters, a good pun or surprising word can pull on our attention in a way which has rhythmic consequences. Even though content is mostly entirely separate from delivery and flow, then, it can clearly influence both in a significant way.

Phonology is interesting in that it is defined by both rhythm and melody, two parameters that are often treated and analysed separately in music analysis, and that belong to two different categories in Edwards’s taxonomy. When the linguistic study of Gilbers et al. (2019) defines flow as containing both rhythm and melody, it might well be because the two are often treated like different aspects of a compound phenomenon in the field of linguistics. In rap, however, melody and rhythm are no longer features belonging solely to the linguistic domain, and the rules of *musical* rhythm and melody are just as, if not more, important than those of phonology. And while melodic features have rhythmic implications in both linguistics and music (again mostly tied to prominence and segmentation, as well as positioning within a structural framework), they are not *primarily* rhythmic. The melodies themselves are important and analysable regardless of the exact rhythms they inhabit. This is not to say that I believe melody can be fully disentangled from rhythm; a melody’s rhythm is a central part of its identity. But the potential exists for considerable variation in the rhythmic structure of a
melody which will still preserve the *melodic structure* (and vice versa). The point is rather that, by considering flow to be *rhythmic* structure, melody can be considered to be part of delivery—it clearly influences the rhythmic experience, but it is much more than just rhythm and should be treated accordingly. This means that one cannot simply define the relationship between flow and delivery as the same as that between structure and expression (because delivery includes melody, which is certainly also structural). Nevertheless, pragmatically speaking, if one limits flow to rhythmic structure, one can think of delivery as that which is neither rhythmic structure nor lyrical/semantic content.

One can also extend the pragmatic delimitation of considering that which is not primarily rhythmic to be delivery rather than flow to other articulative features, like enunciation (or staccato/legato, to use musical terms), dynamic variations and so on. The categories of delivery and flow are entangled and interdependent, but they are also separable and intuitive, and they are a big part of the discourse among practitioners and aficionados. The blurring of the boundaries between delivery and flow, and between what is rhythmic and not, will come up often in this thesis, both in track A6 and throughout the analyses on side B.

**Rhythm: Flow’s direction and expression**

‘What is rhythm?’ might be an even more foundational question than ‘what is flow?’ Again, it threatens to complicate something simple, as tacitly, down to our bones, we all know what rhythm is. Still, to be able to work intimately with the concept of rhythm in an analytical and theoretical manner, we must critically evaluate how we understand, and what we mean when we say or write, ‘rhythm’. In the introduction to *Musical Rhythm in the Age of Digital Reproduction*, Anne Danielsen observes that ‘rhythm comprises an interaction between non-sounding reference structures (…) and sounding rhythmic events’ (2010a, p. 4). This understanding of rhythm as an amalgam of the physical acoustic signal and the way listeners make sense of its structure in their minds and bodies is the natural continuation of a broad tradition of scholarship on musical rhythm.16 Danielsen’s own take on this tradition (2006) is to use concepts from Mikhail Bakhtin and Gilles Deleuze to analyse the interaction between *gesture* and *figure* in classic funk grooves. The gesture is the actual sounding event, while the

---

16 Among others, Danielsen mentions the ‘Uppsala school’ of Ingmar Bengtsson and Alf Gabrielsson and their SYVAR (systematic variation) concept; Tellef Kvifte and Eric Clarke and their respective works on categorical perception (upon which I draw heavily in this thesis); and Peter Desain and Henkjan Honing, Charles Keil, Mats Johansson and Vijay Iyer as scholars subscribing to this way of conceptualising rhythm (Danielsen, 2010a, pp. 4-6).
figure is a virtual reference structure. Perhaps the most important aspect of this distinction is that, according to Deleuze, the virtual is no less real than the actual:

In fact, the virtual is fully real and must be defined as a part of the real object, as though the object resides partly in a virtual domain. (. . .) In parallel, we might conceive of musical reference structures as virtual aspects of the real music, while the sounding events are actual manifestations of the same reality. The music has a part of itself in a virtual domain. (Danielsen, 2006, p. 47)

The topics of reference structures and categorical rhythm perception will be fleshed out extensively in track A3, but for now, the introduction of the virtual and actual—the figure and the gesture—has some important epistemological implications for the entirety of this thesis.

Gesture names a demarcated musical utterance within the fabric of a rhythm. It might be a riff or a vocal phrase, or a part of either, or a group of beats, or just one beat, as long as it is perceived as forming an entity, a sounding gestalt. A musical gesture includes in principle every aspect of this entity—that is, the actual as well as the virtual. Even though one parameter often tends to be the primary characteristic aspect of the gesture—it may be shaped by, for example, timbre, rhythm, or melody—the gesture transcends any traditional division into analytical parameters. (. . .) The figure is a virtual aspect of the gesture and might be conceived of as a proposal or schema for structuring and understanding the gesture. (Danielsen, 2006, pp. 47-48)

Just as the figure and gesture are mutually dependent upon one another—two sides of a single musical coin—this understanding of rhythm sets certain conditions for analysis: first, any analysis must consider not only the physical acoustic signal but also its accompanying reference structures/figure(s), and second, it must proceed with the understanding that there is not necessarily a one-to-one relationship between gestures and figures. ‘At the level of figure, one rhythm might be identical to another, while at the level of gesture, they might be different’, Danielsen continues (2006, p. 50). Likewise, one gesture might have multiple accompanying figures, and some of its most striking effects might only be comprehensible when looking at their congregation. This understanding of rhythm therefore adds an extra layer of complexity to the analysis of sounding music compared to more traditional structural analysis based on scores. It makes it necessary to think through how to represent the object of analysis, and also what kind of truth the analysis can and should produce.
The topics of representation and presentation of music and the cognition of rhythm are central topics of this thesis and will be revisited thoroughly in the next couple of tracks. Related to both, as well as to the word ‘flow’, is another core theme: music’s forward movement. While music is an art form that unfolds as time passes, musical analysis and the act of transforming music from sound to representation or descriptive text is essentially taking this inherently temporal phenomenon out of time.

A music theoretical register (…) tends to place a particular musical work in a kind of contemplative, conditional freeze-frame: a non-space/non-time that allows it to be apprehended outside of its unfolding, as if it is in stasis, present all at once like a painting; this allows one to leisurely examine its different constituent parts, their behaviors and interactions, and the ways these parts articulate with other structures external to the piece. (Daughtry, 2020, p. 7)

While it is impossible to avoid this out-of-time-ness in any printed representation of music (and the emergence of digital representations with moving pictures could be a true paradigm shift in music theory), it is possible to be cognisant of and to emphasise music’s temporal unfolding during analysis and when writing about music. This is a critical argument of the present thesis. The analysis of rap flows, as we will see already in track A2, certainly encompasses many good examples of the out-of-time-ness of the music theoretical register, as such examples focus on global features of the music. Rather than analysing specific successive musical moments, they extract trends and averages to tell us something about the whole of a formal unit of music—a full track or even an entire corpus representing an artist or a genre. While it might be both enjoyable and illuminating to engage with music as an architectonic phenomenon in this way, I am more interested in the ephemeral qualities of music unfolding in time. When engaging with terms like ‘structure’ and ‘framework’ in musical analysis, that is, we must keep in mind that these structures are not actually written in stone but rather constantly shifting and negotiating with one other (some more than others, as we will see later).

With rap analysis having developed into a wide and productive field of research, analysts now have the luxury of focusing on their chosen musical features and aspects while building on the good work of others. To use a linguistic analogy, when we are looking at rap’s rhythmic language, we can consider global features to be its syntax. They tell us what the language’s sentence structure tends to look like. Without syntax, one would be unable to create meaningful sentences, but syntax alone holds very little communicative meaning. The
rhythmic features and techniques analysed in this thesis are more akin to poetic techniques or figurative vocabulary. Or, if we leave the language analogy behind in favour of a musical one, this thesis is focused on ‘licks’ or musical vocabulary rather than on scales or harmonic progressions.

Like my separation of flow and delivery, my epistemological choice of licks over scales is founded on my experiences as a practitioner of rap. When practicing an instrument, playing scales and arpeggios is only one, fairly limited way of developing one’s ability; most musicians also transcribe, copy, rearrange and recontextualise others’ playing as well. While jazz saxophonists fill their proverbial ‘bag of licks’ with bits of Coltrane and Adderley solos, rappers fill theirs with cross-rhythms from Missy Elliott or asymmetrical phrases from E-40. Quite often, this is how I listen to rap flows. My analytical ear is tuned to increasing my rhythmic repertoire rather than necessarily gaining some overarching insight into rap as a genre. Yet, it advances my research nevertheless, as I label, categorise and describe to myself different types of ‘rap licks’. There is also a pedagogical dimension to this. By determining how I would best describe certain features of a rap flow, I begin to see how to reproduce them and, in turn, represent them outside of their immediate musical context.

But how do I choose which licks are worthy of attention? Which rhythmic particularities in a rap flow stand out (or ought to stand out)? Which small parts of the shit make it dope? These are questions of salience or markedness. Both terms describe something standing out from its surroundings, the former typically addressing how and how much it stands out. The latter, on the other hand, is a binary—that which is marked stands out from an otherwise established structure or form. (Something is marked, and salience refers to how it is marked and how marked it is.) I will use ‘marked’ in this thesis. While the concept of markedness originates in the field of phonology (Trubekoj, 1969), it has travelled to many other disciplines. Nigel Fabb, a scholar straddling the fields of literature and linguistics, discusses it in relation to relevance theory: ‘Communication is governed by the principle of relevance, which guarantees that if greater processing effort is demanded (e.g., by a text) then there will be greater cognitive effects’ (Fabb, 2015, p. 267). Something which stands out (that is, demands more cognitive effort) is important in some way, and in the context of rap flows, this importance is aesthetic in nature. Of course, something can be marked in a negative sense—it is wack, not dope—but I will focus on the former in this thesis. That musical enjoyment, or dopeness, is linked to markedness has been explored by, among others, David Huron, whose *Sweet Anticipation* (2006) presents a theory of expectation centred around our emotional
responses to expectation fulfilled or broken. It is most often broken expectations that mark something, but this does not work unless that something is in contrast to the unmarked, which is itself governed by the fulfilment of expectation. There will be more on anticipation, broken expectations and Huron in track A3.

**Context: The scholar, hip-hop and the world**

The cycle of returning to myself as a reference might seem self-indulgent, but there is a method to the madness. While any doctoral thesis will be rooted in the inclinations, preferences and personality of its author, I aim to exploit this quality to my advantage—that is, it is a secondary goal of this thesis to engage with the particular affordances presented by my background and practice as a recording artist and performer. One reason for this is that, though the analysis of rap music has advanced immensely during the years I have been immersing myself in the topic, I often find myself taken aback by some analytical angle, explanatory model or conclusion in the literature. It is rarely anything major or egregious but rather a sentence here or there that makes me think ‘that’s not what it is!’ or ‘it’s the other way around!’ There are, of course, differences regarding what rap is among both scholars and practitioners, but I still think that, in both the layperson’s and the scholar’s discourses, the insights of the performing rapper are not utilised to their full potential. While rappers’ opinions and knowledge might be accommodated at a surface level, their voices are rarely heard in discussions of the nuts and bolts of rap. I believe that there are certain insights to be gleaned from artistic practice which, while not necessarily crucial to finding answers, might aid in asking new and interesting questions.

This latter point is at the core of Donald Schön’s concept of ‘reflection-in-action’, the modus operandi of the ‘reflective practitioner’ (Schön, 1983). In short, the practitioner can access a different type of knowledge and reflection from that of someone on the outside looking in. First of all, there is the tacit knowledge a practitioner accumulates without necessarily being able to (or prompted to) articulate it, and second, there is the practitioner-as-practitioner’s prerogative to encounter and engage in situations which are unavailable to non-practitioners. Or as Schön describes it:

> When the phenomenon at hand eludes the ordinary categories of knowledge-in-practice, presenting itself as unique or unstable, the practitioner may surface and

---

17 A central point for the originator of the term ‘tacit knowledge’ (or ‘tacit knowing’), Michael Polanyi (1958, 1967), is that some knowledge cannot be articulated at all.
criticize his initial understanding of the phenomenon, construct a new description of it, and test the new description by an on-the-spot experiment. Sometimes he arrives at a new theory of the phenomenon by articulating a feeling he has about it. (Schön, 1983, pp. 62-63)

While I will not engage in ‘heuristic research’18 here, and the toolset and theoretical framework I employ will be familiar to most music scholars, Schön’s concept resonates with me. In particular, I am taken with the idea that the practitioner can ask slightly different questions than the non-practitioner and is able to assess existing theories from a different position:19 ‘someone [who] reflects-in-action ( . . .) is not dependent on the categories of established theory and technique, but constructs a new theory of the unique case’ (Schön, 1983, p. 68). I find that when engaging with the musical material from other scholars’ analyses, I often hear and think something slightly different, particularly in terms of what might be significant in the structure of the music.

One of my first thoughts and instincts, then, is typically ‘how can I do this, but differently?’, meaning that what I seek to identify is what in the music moves me, and what that musical feature is. The latter might involve any or all of the following questions: What are the constituents of this musical feature? What other features is it like? What are its potential variants and variations? Why is this feature significant to me? How will this feature be different when recontextualised or altered in various ways? This part-scholar, part-practitioner positioning will permeate this thesis, though the thesis is not a ‘how to rap’ manual. Its descriptions and analyses of music are always based on a listener’s position, not a performer’s.

But I am not only a reflective practitioner in my engagement with this music and culture or the world at large. I am also a (or one could say ‘yet another’) White cis-male academic writing about Black music20, another position which requires reflection, as does the profound

---

18 Heuristic research is a framework developed by Clark Moustakas (1990) which entails ‘stacking the deck’ for tacit knowledge and intuition to instigate ‘eureka moments’ of insight while actively avoiding established patterns of thinking and research. I have dabbled in this approach previously, and while I am not using it in this thesis, I believe it can be very valuable in the study of music, especially for practitioners.

19 One other way of conceptualising this ‘different position’ is via Heidegger’s modes of comportment, where scholars typically rely mainly on erkennen (which could be translated as ‘cognising’), while practitioners might be more adept at umgang (or ‘going-about-the-world’).

20 Note that I capitalise both ‘Black’ and ‘White’ when they refer to a racialized identity. The history and reasoning for the former is well established—see for example The New York Times’s reasoning for using it (Coleman, 2020)—the latter is more contested. I follow Nell Irvin Painter’s (2020) argument in that ‘We should capitalize “White” to situate “Whiteness” within the American ideology of race, within which “Black,” but not “White,” has been hypervisible as a group identity.’
non-Americanness of my specifically Norwegian and more generally European background, as regards both hip-hop culture and academic disciplinary traditions. I will not be able to engage sufficiently with these topics, which would require a thesis of its own, but I will preface the following discussion with a few statements:

1. My work both as a hip-hop scholar and hip-hop artist is cultural appropriation.
2. Hip-hop does in no way need academia, nor does it need my work.
3. This section is not just a simple disclaimer. The problematic aspects of a White man who has come from a particular cultural tradition and who occupies a position of privilege in the racial–colonial hierarchy looking at a Black cultural expression through what Philip Ewell (2021) calls a White racial frame cannot and should not be understated. However, I believe the good my work does outweighs the bad, because anyone engaging with and challenging the inherent Whiteness of the field and society is valuable, even if that engagement, in this case, is done through a thesis which would not exist without White male privilege. While I cannot deny my privilege, I can use it as productively as I am able to in an attempt to make the world a better place.

As stated in the opening paragraph of this thesis, rap, and by extension hip-hop, is a global phenomenon. This globalisation of the cultural expression has naturally led to a variety of local appropriations. Let us compare it to a very different type of cultural expression: pizza. The way pizza has travelled from its roots in Italy (and there is a significant difference between what Neapolitans and what Romans would consider ‘authentic’ pizza) and been appropriated—some would even say subverted—by local traditions like the deep-dish Chicago style or the Swedish variant with bananas and curry powder21 might be frustrating to many Italians, but due to how fantastic pizza is, this appropriation was inevitable: everybody wants pizza, however little they might know about its roots. Likewise, people from all around the world find both pleasure and immense social value in hip-hop. The task for those of us who have been exposed to and fallen in love with an expression of this culture such as rap music while lacking a rounded knowledge of the culture as a whole is to make sure we learn more and are sensitive to the culture in the way we engage with it. Appropriation is not exploitation (even if it clearly can be). Akil The MC of Jurassic 5 wrote the following on the group’s Facebook page:

21 My personal specialty is made with cured fish and lacto-fermented plum (highly inauthentic).
I am a guest in the culture of Hip Hop because I am not from where Hip Hop originated (The Bronx, New York). I’m from Los Angeles. I was invited to this culture in 83 and have been treated as a guest should be treated with honor and respect and nothing but the best. I don’t mind being called a guest or have a problem with someone saying I’m a guest in this culture. I know what I am and not ashamed of it. (Akil The MC, 2021)

Like Akil The MC, I wish to take the stance of being ‘a guest in the culture’. Whether or not my enthrallment with G-Funk music videos in the mid-1990s should be considered an invitation is up for discussion, but I believe it is experienced as such by young, impressionable minds like mine. Of course, my engagement with hip-hop as a musician and a scholar is an appropriation of the cultural expression, but my take on hip-hop (music and research) is derived and delivered with the utmost respect for its roots and history, and I hope and believe that this shines through in my work.

As a call to arms for music researchers to challenge the existing, suppressive structures of ‘music theory’—both the institutions teaching it and the discipline itself—Philip Ewell’s ‘Music Theory’s White Racial Frame’ (2021) sets out to become the most influential publication in music studies in the early 2020s. Painting a picture of the field of music theory, both in general and specifically within US institutions, the article sums up how the field displays and reinforces its approach.

Our white racial frame believes that:

- the music and music theories of white persons represent the best framework for music theory.
- among these white persons, the music and music theories of whites from German-speaking lands of the eighteenth, nineteenth, and early-twentieth centuries represent the pinnacle of music-theoretical thought.
- the institutions and structures of music theory have little or nothing to do with race or whiteness, and that to critically examine race and whiteness in music theory would be unfair or inappropriate.
- the language of ‘diversity’ and the actions it effects will rectify racial disparities, and therefore racial injustices, in music theory. (Ewell, 2021, p. 2)
Taking Ewell’s description of music theory’s problematic beliefs to heart, I see several implications for how I might attempt to combat the White racial frame as a young music scholar. To start with, the first two points are ones I set out to challenge, both through my musical material and through the theoretical and methodological framework I apply to it. The former is obvious—I write about, perform and analyse music rooted in a Black American tradition, not in comparison to other musics but as the central protagonist of its own story and its relation to me, the analyst. The latter is not obvious, as it is by no means clear which applications of which theoretical and/or methodological frameworks would be suitably ‘colour blind’, and throughout the thesis appear applications of frameworks (both theoretical and methodological, with regard to music, poetry, linguistics, psychology or philosophy) which are based in or on the White German-speaking traditions Ewell identifies as foundational for a White racial frame. The very idea of ‘colour blindness’ is, as Ewell notes, problematic in itself, because an important step in understanding and ultimately resolving or ‘de-framing’ the discipline is to both acknowledge and embrace the historical and cultural concepts of ‘Whiteness’ and ‘Blackness’. I have, to the best of my ability, evaluated whether my choices and applications of theories and methods are sensitive to both cultural and musical nuances, and when I have committed to a specific framework—like musical notation or theories on categorical perception—it is because I believe it to be suitable to the analysis of ‘Black music’, however White its origins might be.

Ewell’s third bullet point highlights the institutions and structures of the field of music theory and in particular the perpetuation of the White racial frame by US academic institutions. The picture Ewell paints of the persistent conservative aspects of these institutions is unfamiliar to me as a Norwegian music scholar. Coming from a position of perhaps ignorant naïveté, I have always reacted with surprise and bemusement when reading or being told about how US music theory programs have obligatory German language or Schenkerian analysis classes. The former would never be required in a Norwegian institution, and the latter is only briefly mentioned in introductory classes as a marginal practice at best. From the outside looking in, it seems that I am in yet another privileged position in that I have never been discouraged from pursuing an interest in ‘popular music’ (an antiquated and loaded term if ever there was one) in general or hip-hop specifically. While there remain clear challenges concerning both inclusivity and structural Whiteness in Norwegian institutions as well (speaking for myself, the music education I received some fifteen to twenty years ago was heavily weighted towards European ‘art music’), I take pride in belonging to a group and lineage of
(admittedly, mostly White) music scholars who actively advocate ‘for a restructuring of our racialized structures—a deframing and reframing of the white racial frame’ (Ewell, 2021, p. 1). Happily, that is, Ewell is most definitely not a lone voice but part of a large progressive movement of music theorists whose voices—both in publications and in the social media discourse—are at once inspirational and encouraging.

However, as I am a White, privileged man, my engagement with hip-hop and rap both as an artist and a scholar is, again, fundamentally an act of cultural appropriation, and the Blackness of the cultural origins of the music I make and analyse will always be contrasted by the Whiteness of my cultural background. If some believe that my music or academic work is less valuable because it is less Black than the music inspiring it, that is their prerogative, but the type of cultural appropriation that this thesis and my music exemplifies is not one of Whiteness attempting to subjugate Blackness but rather one celebrating Black music in general and hip-hop and rap specifically. White people engaging with Black music cannot stop the music from being Black—there is no requirement of purity, as Imani Perry argues in her influential Prophets of the Hood: Politics and Poetics in Hip Hop: ‘To deem something French or English rarely implies that there were no Germanic cultural influences, or Irish, or even Algerian’ (2004, p. 11). Hip-hop has taken over the world, and it will stay Black even when it is also Norwegian, White and middle-class.

Hip-hop’s history in Norway is older than me, and its story is chronicled by journalist Øyvind Holen in his two Hiphop-hoder books, the first (2004) subtitled Fra Beat Street til bygderap [From Beat Street to small town rap] and the second (Holen, 2018) subtitled Hvordan Karpe Diem og generasjon 1984 tok en undergrunnskultur til pophimmelen—og endret Norge på veien [How Karpe Diem and generation 1984 brought an underground culture to pop heaven—and changed Norway on the way]. As the subtitle of the first book indicates, hip-hop first struck Norway via the legendary motion picture Beat Street in the summer of 1984, inspiring the first generation of Norwegian hip-hoppers. While the impact of Beat Street and hip-hop upon Norwegian society at large was minimal at the time, certain early pioneers fell in love with the culture and cultivated it in private and in small underground scenes.

22 Karpe Diem (now known simply as Karpe) is the most commercially successful Norwegian rap act of all time. Rappers Chirag (Chirag Rashmikant Patel) and Magdi (Magdi Omar Ytreeide Abdelmaguid) have consistently sold out the country’s largest venues, headlined the largest festivals and been the focal point of political and cultural discourse in Norwegian mainstream media since the turn of the millennium.
In the 80s, breakdancing was compared to the yo-yo, a fashion thing that would die out fast. In many ways, that was correct, but at the same time, there was a deeper culture in the background. The few of us that dove deep kept going, and the culture has survived. Today hip-hop is finally recognised as a culture. (Bjørn ‘Sean’ Hagen to Øyvind Holen in Holen, 2004, p. 24, my translation)

Put simply, the Norwegian hip-hop scene at the start was marginal, underground and hung up on concepts like authenticity and ‘realness’. The overly commercialised appropriation of rap by the mainstream media (particularly prominent in TV commercials) was the epitome of what hip-hop was not to the Beat Street generation. At the same time, Norwegian hip-hop was inevitably very White. Outside of a very few notable exceptions (like Leo Fossen of L.A.W., A-Team and B.O.L.T. Warhead), Norwegian hip-hoppers were White, and with Norway’s welfare state as a safety net, the working class was and is relatively affluent compared to its counterpart in the United States.23 It took a long time for rapping in Norwegian to be accepted in the quite hip-hop-culturally conservative hip-hop scene in Oslo,24 but when it finally broke through, it heralded a new era, spearheaded by the younger ‘generation 1984’ described by Holen as

a generation that falls for hip-hop in the early 1990s, at a time when the influence of New York–based breakdance, DJ-ing and graffiti is fading, supplanted by songs about ‘pussy, weed and alcohol’ from gangsta rappers in California and the South. (Holen, 2018, p. 11, my translation)

This generation, while perhaps less ‘real’ according to the Beat Street generation’s hip-hop ethos, is also more culturally diverse, both because more and more Norwegians at the time had immigrant backgrounds and because the scene was less Oslo-centric. This development of a diversity of expression, sound and cultural background has perhaps mirrored both hip-hop’s development in general and Norwegian society’s development in particular. At the time of this writing, the Norwegian rap scene is dominated by artists with compound cultural backgrounds,25 including the aforementioned Karpe (Magdi has one Egyptian and one

---

23 However, it must be noted that hip-hop in Norway has historically had strong ties to the political far left and has been strongly lower middle class and also associated with ‘outsider’ narratives.
24 The paradox of ‘being real’ entailing that one would have to rap in a foreign language is both odd and common to most non-English-speaking countries. Holen (2004, p. 167) attributes Norway’s relative lateness in embracing mother-tongue rapping to the tøyserap (meaning something like ‘messing-around rap’) in mainstream radio and television.
25 I avoid using the term ‘diversity’ here both to follow Ewell’s advice about challenging the White racial frame and to emphasise the essential Norwegianness of these artists. Without going into a discussion of the implications of the term bindestreksnordmann or ‘hyphen-Norwegian’, the mentioned artists are often referred to
Norwegian parent, Chirag’s parents are Indian, and both grew up in Oslo), Arif (born in Copenhagen to parents from Zanzibar, grew up in Norway), Kamelen (Norwegian-American mother and an absent father from Botswana; refers to himself as afroamerikansk bergenser)\(^{26}\) and Musti (Somali parents). While Norway and the United States are immensely different, the Norwegian adoption and appropriation of hip-hop has done much to conserve and celebrate its founding spirit.

This thesis is not about hip-hop culture, though, but about rap. And if the academic pursuit of rap’s aesthetic intricacies is to be successful, the scholar studying it must be intimately familiar with this music. There is (as we will see in track 2) an impressive array of published analyses of rap, and most of the time these scholars’ enthusiasm for hip-hop shines through. Still, there are also varying degrees of familiarity with the music and culture, and, as mentioned, there are few practitioners lending their voices to the choir. Again, I believe the positioning of the ‘reflective practitioner’ is likely to be fruitful in this field, because a hands-on engagement with the music can produce insights which might elude others. Eileen Southern writes in *The Music of Black Americans* that ‘serious study of African-American music requires getting to know the music, which means listening to it and, if possible, performing it’ (1997, p. xx). While I might be only an uninvited (or unwittingly invited) guest in hip-hop culture, I have—with sincere enthusiasm—listened to, made and performed rap music for many years. I am not a conventional ‘insider’ in the main branch of hip-hop’s historical culture, but I can claim the status in one grafted-on twig,\(^{27}\) and I have, as Southern requests, both listened to and performed rap—and gotten to know some version of it quite intimately. Of course, Southern’s ground rule for music analysis is by no means exclusive to the study of Black music. However, as a younger subdiscipline in a well-established academic field, rap analysis must critically evaluate which methodological and theoretical approaches its adherents take, and whether those approaches are indeed suitable for analysing the music at hand or instead simply those most convenient to the trained music scholar. Like virtually everyone else in the privileged position of being funded to write a lengthy treatise on music, I

---

\(^{26}\) See Holen (2018, p. 194). Kamelen’s identity as a bergenser (from Bergen) is also significant. In Norway, no other city has the same idiosyncratic identity tied to it. Likewise, particularly in the hip-hop scene, the part of Bergen from which Kamelen hails, Loddefjord, has its own strong identity as a drabantby (a less affluent satellite town with high-rise developments) where the youth are often instigators of cultural movements, particularly in the arts.

\(^{27}\) The inspiration for positioning my ‘insiderness’ on some sort of insider–outsider spectrum comes from Chris Stover’s positioning through self-identification and background in his dissertation (Stover, 2009, pp. 32-33)
have been trained in a discipline that ‘locates [certain] Western European and Euro-diasporic practices as an unmarked norm’ (Stover, 2022 [In press], p. 2) meaning that my toolbox includes some insurmountable biases towards which musical features and structures are readily identifiable and representable. It is impossible to separate myself and my practice from these biases, but I attempt to be conscious of them and how they impact my work, and I am careful to make the reader aware of them as well.

Ahead of the analyses of side B of this thesis, I sum up my methodological approach in a short sample analysis in track A6, but the question of what a suitable toolbox for rap analysis should be informs the entirety of side A.

**Analysis: What we do, and how do we do it**

While it could be defined as straightforwardly as pointing out something dope in a piece of music, the tradition of what musical analysis is—or, perhaps more pressing, what musical analysis should be—has fuelled a polemical epistemological debate throughout the twentieth century (and beyond). One of the strongest voices in this debate is that of the late musicologist Joseph Kerman, whose *Contemplating Music: Challenges to Musicology* (1985) represents a foundational critique of what he considers to be an excessively positivistic lop-sidedness to the field of musicology. In particular, Kerman called for musicology to ‘to confront the work of art in its proper aesthetic terms’ (Kerman, 1965, p. 65) and embrace the fact that all musical analysis is in fact musical criticism: ‘the study of the meaning and value of art works’ (Kerman, 1985, p. 16). He warned against analysis done in the interest of illuminating a tenet of theory, although ‘one cannot always draw sharp distinctions. An incidental demonstration in a theory treatise may illuminate a work of art, just as a critical essay may incidentally illuminate a point of theory’ (Kerman, 1985, p. 69). Kerman’s issue was that the analysis of the constituents, form or structure—that is, the contents—of the work of art had become the primary manner of engaging with music for musicologists, at the profound expense of a more holistic approach: Kerman’s idealised ‘criticism’.

This very sharp stance met with some opposition. Another of the most influential voices in the last century’s musicology, Leo Treitler—like Kerman, a representative of what became the branch of the discipline dubbed ‘New Musicology’—questioned whether Kerman’s dichotomy between positivism and criticism is an accurate rendering of how musical analysis is practiced and what its goals are:
The very idea of such a dualism bears within itself the premise that the products of ‘positivist musicology’ do realize an intention to embody positive knowledge unaffected by subjectivity. Ironically, this polarized portrayal of the contemplation of music conveys an inflated respect for the power of positivist thinking in its own sphere, and a low estimation of the possibilities of subjectivity outside of its own, different sphere (criticism). (Treitler, 1989, p. 376)

Treitler stressed that there is actually no vast gulf between ‘thinking about texts’ and ‘establishing texts’, and that the latter is not necessarily mere positivism or formalism but also in many ways criticism. Regardless of the validity of Kerman’s diagnosis of the state of musicology in the 1980s, however, he sparked a sprawling debate about what music analysis could and should aim to achieve.

An interesting rebuttal to Kerman’s discussion of the relationship between analysis and criticism is the aptly named answer to Kerman’s 1980 article ‘How We Got into Analysis, and How to Get Out’: Kofi Agawu’s ‘How We Got Out of Analysis, and How to Get Back In Again’ (2004). Revisiting the discourse within the field of musicology—or perhaps rather the fields of New Musicology, music theory and ethnomusicology—with all the nuance of hindsight, Agawu set out ‘to restore a vision of music analysis that stresses its affinities with performance and composition without denying its autonomy’ (2004, p. 279). Agawu argued for the value of analysis as it is connected to musical practice, both performance and composition, and the value of engaging with music through analysis in itself.

The aim of the 50th analysis of the ‘Appassionata’ is not to add incrementally to the previous 49 (although it can do that); it is rather to provide the analyst with an opportunity to make the ‘Appassionata’ his or her own. Just as we do not ask of the 50th recording of the ‘Eroica’ what new knowledge it adds to the previous 49, so we should not expect of an analysis that it add to some existing body of positive knowledge. (. . .) I believe that the more fundamental motivation lies in the desire to inhabit temporarily a certain sonic world—and to enjoy the sensuous pleasure of so doing. Only in response to certain institutional imperatives in the modern

---

28 Treitler here pointed to Kerman’s insistence that ‘musicologists as a corps spend so much more time in establishing texts than in thinking about the texts so established’ (Kerman, 1985, p. 226), which is admittedly something of an overly simplistic strawman.
academy/economy does the motivation for analysis become primarily positivistic in nature. (Agawu, 2004, pp. 274-275)

The idea of music analysis being primarily a pleasurable act meant ‘to overwhelm, entertain, amuse, challenge, move, enable, indeed to explore the entire range of emotions, if not in actuality then very definitely in simulated form, at a second level of articulation, so to speak’ (Agawu, 2004, p. 280), and only secondarily meant to make explanatory or epistemological points, is an attractive one for a musical analyst. And if it were true, as Agawu argued, that ‘analytical knowledge resists or escapes verbal summary’ (2004, p. 274), one might even conclude that performing an analysis (or at least engaging with the music in an analytical manner, which is really what analysis is in its simplest form) is the only way to obtain some of the many insights which are available about a piece of music.

If part of the value of analysis is the analyst’s subjective experience of ‘making the music their own’, then the perspectives, inclinations, goals, values—the very identity of the analyst—is critical to the analysis itself and whatever epistemological points it might make. I have already implied the same here. So: What are my goals and motivations going into my own analyses, outside of the obvious need to respond to my research questions and thesis title? And what do they mean for the nature of the analyses?

Following Agawu’s notion of analysis as performance, one central aspect of my process is the pleasurable and educational experience of sharing an analytical space with music I already enjoy so much. By tapping into my positioning as a reflective practitioner, I will attempt to accomplish what J. Griffith Rollefson calls for in Flip the Script: European Hip Hop and the Politics of Postcoloniality (2017, p. 10): namely, ‘hip hop close readings in both form and content’ which are not the result of mere scholarly work on hip-hop or even the scholarly work of a rapper. Instead, I intend to ‘move past [my] subject position as “scholar of hip hop” and truly take on the mantle of “hip hop scholar” by doing scholarly work in a hip-hop way’. The way this is most visible in my work is on the surface, where the language, discourse and graphics employed are ‘hip-hop flavoured’, as in the thesis’s title, the hand-drawn figures, and the personal anecdotes contextualising my analytical and theoretical discussions. Most importantly to me, doing my scholarly work in a hip-hop way means approaching analysis and the act of doing music theory like I approach doing hip-hop—as a creative, playful and enjoyable act of immersing myself in a musical and cultural expression which I consider to be

I might even argue that the hip-hop flavour permeates the style of my prose as well.
radical, democratic, socially conscious and capable of being a positive transformative force at both an individual and a collective level. I want my analysis to reflect the joy of my approach or mindset and even to be enjoyable (or even inspirational) for the reader.

Another aspect of my process is the more traditional ‘woodshedding’ of the musical practitioner. Transcribing, deconstructing and recontextualising musical ideas are common ways of expanding one’s own musical vocabulary, and this is how I initially started doing rap analysis. Zooming in on various constituents of the musical whole certainly supplies insight into techniques and approaches, if not into the musical whole itself. It might be the type of formalist ‘analysis illuminating a piece of theory’ which Kerman would condemn, but, contrary to the positivist tradition, I make no claim about my truths as either objective or universal. In fact, my claim is the exact opposite—the subjective position of both the analyst and all of the analyst’s analyses is central to what makes the analysis valuable.

Finally, the third aspect of my process is a theoretical and discursive one. In both academic and colloquial discourse there is perhaps not a ‘need for’ but at the very least a gap in the naming of musical techniques, their variants and their role in rap music as a whole. My analyses aim to highlight, present and explain various musical, poetic and theoretical concepts which I believe are significant to rap. This may be useful both to the field of rap analysis and to music makers, teachers and students of both the theoretical and practical aspects of rapping, as well as aficionados and ‘heads’ in general.

All in all, my subjective position and approach are core to my analyses, as they are for everyone else who immerses themselves or dabbles in rap analysis. The next track is dedicated to those others who came before me, creating and populating the field of rap analysis.
Track A2: Rap analysis

As the field of hip-hop studies has emerged and the topics of hip-hop and rap have found their way into a range of disciplines and academic traditions, subfields have branched off of the original trunk as well. The branch which sprouted this thesis has grown large and strong over the past three decades or so: rap analysis. In this track, I will present the principal areas of research which have focussed on flow and other non-semantic parameters of rap (that is, I will not present the body of work concerned with lyrical content). Furthermore, I will divide this rap research into categories based upon the features and/or parameters they address, unpacking and otherwise explaining how rap flows are analysed and (re-)presented, how the different approaches relate to mine, and how I will engage with them. This track contributes to my response to the second main research question—How have rap flows been analysed?—and specifically the follow-up questions Which tools have been used? and What are the tools’ underlying theoretical assumptions?

The early days of rap analysis

The first published scholarly musical-technical analysis of a rap flow is Robert Walser’s ‘Rhythm, Rhyme, and Rhetoric in the Music of Public Enemy’ (1995), where transcriptions of both the musical background (that is, ‘the beat’) and the rap are presented in traditional music notation, analysed and interpreted. Particular important are Walser’s musings on polyrhythm, groove and repetition, and the importance of those aspects to rap music in general. They are among the defining musical features of what we consider groove-based music, and Walser was writing at a point in time where there was still a dire need for a ‘defence’ of rap’s status as a legitimate musical expression. Walser denounces some artists’ and journalists’ position that rap’s lack of focus on melody and harmony (as these parameters were once typically analysed) makes it less worthy. His approach to its analysis, on the other hand, clearly gestures to New Musicology, and he engages in an extensive interpretation of the parallels between the musical features, lyrical content and extramusical messaging found in both Public Enemy’s explicit political engagement and the societal context of ‘Fight the Power’ when it was released.

A central work in the academic history of rap analysis is the already referenced Rap Music and the Poetics of Identity by Adam Krims (2000). This book works to reconcile the often

---

30 Walser does build on some previous ethnographical and musicological explorations of rap, perhaps most notably Cheryl Keyes’s doctoral thesis ‘Rappin to the Beat: Rap Music as Street Culture among African Americans’ (1991).
competing approaches of ‘music theory’ and ‘musicology’ and describes and applies ‘musical poetics’ as ‘a subset of music theory that addresses the organization of sound as part of broader cultural processes’ (Krim, 2000, p. 14). For the more technically focused body of work owing its heritage to Krim, chapter 2 of Rap Music and the Poetics of Identity has left an obvious legacy. Krim outlines an extensive taxonomy of different rap genres, as well as the different ‘flow styles’ I discussed in track A1. His approach to the visual representation of the music in his analyses has also left its mark on the field. Eschewing traditional notation in favour of a grid of ‘beat classes’ (equivalent to sixteenth notes and expressed as ‘0/0 1 2 3 1/0 1 2 3 2/0 1 2 3 3/0 1 2 3’ in his analyses of flow and ‘1xyz2xyz3xyz4xyz’ in his analyses of the musical background), he presents an economical visual representation focussed on rap’s rhythmic constituents rather than melody and/or harmony.

The explosion of rap analysis

After Rap Music and the Poetics of Identity, the field of rap analysis evolved in many directions via theoretical and methodological approaches rooted in various disciplines and flavours of music research. Most of these academic publications have been journal articles, book chapters in monographs, or anthologies with a broader hip-hop focus such as Felicia Miyakawa’s Five Percenter Rap: God Hop’s Music, Message, and Black Muslim Mission (2005) and the Justin Williams–edited The Cambridge Companion to Hip-Hop (2015). In fact, until Mitch Ohriner’s Flow: The Rhythmic Voice in Rap Music (2019b), no academic books had focussed entirely on the technical analysis of rap. In the decades between Krim and Ohriner, that is, rap analysis in book format was found in such popular scientific publications as the aforementioned How to Rap: The Art and Science of the Hip-Hop MC by Edwards (2009) and its sequel, How to Rap 2: Advanced Flow and Delivery Techniques (2013). Some doctoral dissertations have been dedicated to rap analysis as well, both fully—including those of Nathaniel Condit-Schultz (2016b) and Robert Komaniecki (2019)—and partially (Horn, 2010). Another book publication not fully aligned with either of the above categories is Martin Connor’s The Musical Artistry of Rap (2018), which consists of very technical transcriptions of rap flows and a heavily academic discourse which was not peer reviewed.

31 When representing the different layers of the musical background, Krim uses the names of the notes and chords rather than traditional notation, placing these into the 1xyz2xyz3xyz4xyz grid.
32 Including, notably, many contributions to the online journal Music Theory Online, which take advantage of the possibility of using audio examples.
Rather than presenting all the publications that have contributed to the field of rap analysis in chronological order, I will group them according to the features of rap on which they focus while comparing their respective approaches. Some scholars, such as Mitch Ohriner and Kyle Adams, have published multiple works which belong in different categories, and the larger monographs like Ohriner’s *Flow* and Komaniecki’s dissertation naturally span a range of topics, so their work will be found under multiple subheadings throughout this track.

**Corpus studies and a rap flow’s *global features***

One popular approach to analysing rap involves the *corpus study*, where the analyst extracts information about rap from a collection of measurable data from a corpus of music. As is often the case for statistical analysis of large sets of data, rap-analysis corpus studies tend to focus on identifying trends and establishing common patterns within the corpus. The corpora and the selected data can be framed as representative of all rap music (or all music, even) or framed much more narrowly. For example, the corpus analysed in Horn (2010) is limited to a single artist (Snoop Dogg), and in Duinker and Martin (2017) the focus is on a specific era of hip-hop music (the so-called golden era from the mid- to late 1980s to the mid- to late 1990s; the corpus in the article uses 1986–96). For the corpora seeking to represent the entire genre, the question of sampling is paramount, and this was the topic of several efforts in a single issue of *Empirical Musicology Review* in 2016—articles by Ohriner (2016a) and Condit-Schultz (2016c) were accompanied by three shorter commentaries, with Ohriner commenting on Condit-Schultz’s article (Ohriner, 2016b), Condit-Schultz commenting on Ohriner’s (Condit-Schultz, 2016a) and Jakob Gran commenting on and comparing the two (Gran, 2016). The choice of which parameters to measure and how to measure them is not necessarily straightforward either, which is an interesting point of contention within the discourse. One example of this contention is Jonah Katz’s commentary on Duinker and Martin’s study (Katz, 2017), where he recognises the study’s merits and innovativeness but expresses concern about some of the measured parameters (such as a subjective encoding of the rappers’ enunciation or ‘sharpness’ of consonants). Since quite a few corpus studies of rap have been published, such nuanced discussions among their authors are useful when it comes to evaluating whether the results can be aggregated across studies in any meaningful way.

Regarding the specific application of data which would be involved in analysing rap flows, corpus studies have established evidence for many aspects of rap’s common form and structure, as well as the development of those elements throughout the history of hip-hop. This type of application (representing, as it does, a foundation for further research and data for
comparison purposes) is the stated goal of Condit-Schultz’s dissertation project (2016b) ‘MCFlow’, summarised in article form in the aforementioned contribution to the *Empirical Musicology Review* (Condit-Schultz, 2016c). Here, the gradual formalisation of rap’s form is presented through numbers and graphs. Among the dissertation’s significant findings is its proposed standardisation of a rap song’s form: rap songs trended towards having two or three verses spanning sixteen bars of music each, and they deviated less and less from this standard once hip-hop really started to overtake mainstream radio and television around 1995 or so. Alongside the standardisation of the song form, other trends also manifested, so Condit-Schultz is able to show that the average tempo of a rap track has steadily decreased even as the *rhyme density* and *rhyme chain length* have steadily increased (these concepts are revisited in track A6). The corpus also clearly indicates that the most common placement of both rhymes and phrase endings\(^{33}\) is on or around the fourth beat of the ubiquitous 4/4 musical metre (other musical metrical frameworks in hip-hop do exist, but they are incredibly rare compared to 4/4).\(^{34}\) Condit-Schultz also showcases how his data can be used to compare different rappers’ (or tracks’ or verses’) styles and inclinations, including, for example, *rhyme entropy*, or the variation in the distribution of rhyme positioning within the metre. High entropy means a great deal of variation in the rhymes’ musical metrical positions, while low entropy means that most of the rhymes are placed at one particular metrical position (that is, on the fourth beat).

Using corpus analysis to say something about a specific rapper or a specific flow or verse is the topic of Ohriner’s article in that same issue of *Empirical Musicology Review* (Ohriner, 2016a). His corpus analysis reveals the same tendencies regarding typical rhyme position and segmentation (Ohriner uses ‘line ending’ as a measured parameter, a term to which I will return later in the thesis) as Condit-Schultz’s analysis—that is, rhymes usually appear on or around the fourth beat of a bar and align with the end of a segment of performed text (a ‘line’ or ‘phrase’). However, Ohriner applies this knowledge to highlighting a particular rapper’s deviation from the typical structuring of the flow when encountering a ‘beat’ (as in musical

\(^{33}\) The encoding of ‘phrases’ in ‘MCFlow’ is necessarily somewhat rudimentary, Condit-Schultz writes: ‘This approach is very simplistic, doing little justice to the complexity of phrasing in rap flow. MCFlow contains a wealth of other prosodic, and syntactic, information which could potentially enable much more complex, nuanced, analysis of rap phrasing’ (Condit-Schultz, 2016c, p. 140). The challenging nature of concepts like ‘phrase’ in rap flows will be explored extensively later in this thesis, but however simplistic Condit-Schultz’s phrase-data analysis is, the value of the consistent measurement of something which is at least related to the segmentation of rap flows into smaller units should not be underestimated.

\(^{34}\) Rap scholars unnecessarily belabour this point and are equally fascinated by the very few exceptions to the rule as well—see, for example, Komaniecki (2019, pp. 37-41) and the present thesis (in track B4).
background) which affords different musical metrical interpretations. The work in question—OutKast’s 1996 track ‘Mainstream’, featuring T-Mo Goodie—is unusual in the sense that while the drum track is the traditional duple ‘boom bap’ with a snare drum backbeat and kick on beats 1 and 3, there are several other elements in the track which suggest a 3/4 metrical framework. The guitar part, which is the main melodic component of the track, ‘would undoubtably be understood as triple meter material (…) where it heard alone’, Ohriner observes (Ohriner, 2016a, p. 155). The hypermetric organisation also points to groupings of three bars if the basic time signature is 4/4 due to how the musical material is looped (which would result in the more common four-bar grouping if the time signature were 3/4), and the track’s chorus (or ‘hook’, as Ohriner refers to it) also reinforces this inclination, as it is much more symmetrically organised if interpreted as four bars of 3/4 than three bars of 4/4. Ohriner then shows how the different rappers on the track emphasise different metrical organisations in the ways they place rhymes and phrase/line endings, with T-Mo Goodie emphasising a triple metre which stands out in relation to the analysed corpus.

In all, the corpus studies of Condit-Schultz and Ohriner, plus a third corpus study with a different (linguistic) disciplinary background—Jonah Katz’s ‘Hip-Hop Rhymes Reiterate Phonological Typology’ (2015a)—show that rap flows have certain structural tendencies, and that using variations which go against these conventions is a strategy rappers employ in their flows. Rather than deriving these insights from specific local examples of this practice, however, corpus and corpus-assisted studies express these tendencies as global trends averaged from a corpus.

**Rhythmic surface structure: different schools of representation**

The raw data employed in corpus studies are extracted from transcriptions of rap flows, and such extraction of global features and statistical analysis demand consistent encoding of various parameters across all transcriptions. The majority of rap flow analysis does not concern itself with global features, however, and in most of the analyses which do employ transcriptions of some kind, the transcriptions are adapted to the specific analysis at hand rather than to any sort of global standard. When analysts zoom in to look at specific musical passages, then, they often employ transcriptions of the rhythmic surface structure of the rap flow. While these serve to highlight similar musical features and rhythmic categories, they vary considerably in both the visual presentation of the transcriptions and the transcriptions’ analytical goal(s). There are, however, two main ‘schools of transcription’ in rap flow
analysis—the grid-based and the notation-based—mirroring the two early works (by Krim and Walser) presented above.

While some researchers, like Miyakawa (2005, Chapter 4), have used Krim’s grid-based system with only minor adaptations, others have developed intricate systems featuring the graphic coding of various features. One of the most prolific rap analysts out there, Kyle Adams (2008, 2009, 2015, 2020), is a proponent of a spreadsheet-style graphic layout with one row per bar of music and columns for the beats and their subdivisions (the sixteenth note is the standard subdivision applied, but Adams adapts the grid to thirty-second notes and triplets where necessary). The syllables of the rap flow are placed within this grid and highlighted graphically depending on the features which are most relevant to the analysis. Adams marks stressed syllables with bold type and codes different rhyme classes with different colours. The rhythmic patterns created by these stresses and rhymes are the material on which Adams bases his analyses. In his article ‘Aspects of the Music/Text Relationship in Rap’ (2008), the topic is ‘how rappers incorporate rhythms, groupings, and motives from the underlying music into the rhythm of the lyrics’, and this approach and system have been adopted and adapted by other scholars as well. Rowan Oliver, for instance, uses the spreadsheet-grid system to analyse how Rakim’s flow on the track ‘I Know You Got Soul’ (1987) lines up with the composite drum break (that is, a chopped-up sample) of the musical background (Oliver, 2015, pp. 136-148). In Adams’s second article employing the same system, ‘On the Metrical Techniques of Flow in Rap Music’ (2009), the relationship between flow and musical background has ceased to be the topic. Instead, Adams presents what he considers to be the techniques which constitute flow. While I introduced his suggested ‘articulative techniques’ in track A1, the article mainly revolves around exemplifying his ‘metrical techniques’: the placement of respectively rhyming and accented syllables, the degree of correspondence between syntactic units and musical metrical units (bars), and the number of syllables per beat. All these features are clearly visible and measurable using the grid-based system.

Multiple different grid-based systems have in fact been employed in the field of rap analysis. Horn (2010) adapts the traditional metrical grid of the field of linguistics and marks the hierarchical layers of prominence using stacked crosses. This prominence grid represents the musical metrical framework rather than the prosody, however, so the resulting grid itself is

35 A ‘rhyme class’ is a description of what connects a group of rhyme instances. If two words are connected to each other through two rhyming vowels ‘I’ and ‘A’, they belong to the ‘I-A’ rhyme class.
not much different from Adams’s grid. Other grids are graphically dissimilar to those employed by Adams and Horn but aim to represent the same types of features. Jeremy Page (2018) plots different layers of linguistic constituents along a circular grid. This ‘flowprint’ reads like a clock’s face, with the full verse subdivided into bars, beats and sixteenth-note subdivisions, and colour-coded layers representing (from the smallest to the highest hierarchical unit) phones, syllable nuclei and stressed syllables. Thus, Page represents stressed syllables like Adams but uses a separate layer rather than bold type; rather than plotting rhymes, however, he employs consistent colour coding of all phones across flowprints so that matching syllables (that is, those with the same combination of phones) will be the same colour. Ohriner develops yet another type of grid, first in his aforementioned corpus-study article (2016a) and then refined into the system used in his book (2019b). Page dubs Ohriner’s grid ‘the molecular method (…) due to its visual similarity with common representations of molecules’ (Page, 2018, p. 3) and praises its precision and visual clarity. Ohriner’s system employs a single horizontal line with vertical lines showing the musical metrical hierarchy. The lyrics are placed below the line (as though the line were a single staff line of traditional notation), and numbered black circles indicating rhyme are placed above (different numbers indicate different rhyme classes, while rhymes spanning more than one syllable have their numbered circles connected). On the timeline, each syllable is marked by a ‘notehead’ of sorts—large circles for accented syllables, small circles for unaccented syllables, and triangles for ‘manually corrected accents’. The reason why the latter must be part of the system is because Ohriner has created a semi-automated system wherein a computer algorithm assists in annotating the stressed syllables in the transcribed lyrics, but this algorithm is not flawless, and some manual input is required to achieve a completely satisfactory result. In general, Ohriner is a proponent of the application of CMA methods (‘computational music analysis’) and the idea of using ‘reproducible research” to assist his hermeneutics-based close readings of musical analysis (Ohriner, 2019b, p. xxxiii); his semi-automated system of transcribing and representing the surface structure of flows is an integral part of this project. A significant theoretical claim Ohriner derives from his grid-based system is that rap flows can generally be placed into ‘vocal groove classes’ consisting of various inter-accent intervals (IAIs) which potentially start at different points in the groove class’s

---

36 ‘Phones’ is a term for the smallest linguistic segments.
37 Ohriner’s methodology is both extensively and eloquently documented in the first part of ‘Flow…’ (2019b), and while it will only be briefly discussed here, it is well worth a deep-dive for those interested in the (computer-assisted) representation of rap flows.
rotation. For example, Eminem’s ‘Lose Yourself’ (2002) belongs to the groove class [233233], starting at position 2 (Ohriner, 2019b, p. 83).

The grid-based systems presented thus far all use a standard resolution of sixteen vertical rhythmic units per horizontal line, which then corresponds to a bar of music. This is not the only possible resolution, however, as can be seen in some publications with a slightly different analytical focus. The ‘flow diagrams’ employed by Edwards (2009, 2013) and others (which will be scrutinised in track A4) use a resolution of one column per beat, which makes sense as Edwards’s focus is not on rhythmic patterns and motifs but on rhyme schemes and larger structural units like bars and lyrical segments (phrases/lines). Likewise but in the opposite direction, Gilbers et al. (2019) use different visualisations of rhythmic subdivisions, including a resolution of sixty-four grids in one of them and ‘a temporal resolution of 256th notes’ (!) in another. The latter uses a visual grid of sixteen boxes (representing sixteenth notes), however, with the greater resolution represented by a line within each sixteenth-note box. This approach seeks to identify and represent microtiming, a topic to which I will soon return, but its representational tools are clearly reminiscent of the lower-resolution grid systems presented above.

The other main system of representing rhythmic surface structure in rap flows is more flexible by design when it comes to representing subdivisional categories. And while grid-based systems have established themselves as a strong contender to the title of ‘most used representational system for rap flows’, the stalwart in the history of music analysis—traditional notation—also appears regularly. Amongst the most prolific rap analysts who prefer traditional notation (although he dabbles with the sixteenth-note grid system on occasion) is Oliver Kautny. In the German publication ‘Ridin’ the Beat: Annäherungen an das Phänomen Flow’ (2009), he transcribes both flow and musical background using traditional notation which is slightly adapted to the demands of rap analysis. Kautny uses a single staff line and treats the rap as a percussion instrument, but he does accommodate up to four different pitch levels to represent the constrained melodic content of the transcribed musical segments. Like most analysts, he also uses colour-coding to indicate rhyme. Kautny uses the same system (and some of the same examples) in ‘Lyrics and Flow in Rap Music” (Kautny, 2015), his contribution to the anthology The Cambridge Companion to Hip-Hop (ed. Williams, 2015), to showcase how rhymes and stress patterns in the syllables of the flow create polyrhythms and crossrhythms.
Similarly, most of the analysis in Robert Komaniecki’s doctoral thesis (2019) uses the same approach of traditional notation with colour outlines for rhymes. The thesis is divided into three main parts focussed on the rhythmic and metrical parameters and features of flow, rhyme and pitch, respectively. Komaniecki introduces several concepts which might be familiar to avid hip-hop heads but not so much to academics, such as the flow type ‘stutter rap’ and the rhyme technique ‘rhyme shift’. The analyses of ‘rhythmic motives’ and ‘rhythmic cells’, which explore the rappers create various parallelisms at both regular and irregular intervals to both emphasise and contradict the metrical and hypermetrical frameworks, are particularly successful examples of the use of traditional notation to analyse rhythmic surface structure. Komaniecki’s transcription-focused approach is also the closest match to my own initial foray into rap analysis, my master’s thesis ‘Betre flows enn Akerselva . . . ’ (Oddekalv, 2017), where pages densely packed with figures in traditional notation present various flow techniques and the different flow styles and approaches of the Norwegian rappers Elling Borgersrud, Runar Gudnason and Lars Vaular.

**Microrhythm, ‘expressive timing’ and (dis-)alignment between rhythmic layers**

A significant area of rap research involves the temporal minutiae—rhythmic units which are smaller than those into which we typically divide musical time, and/or the small variations in duration between categorical rhythmic units. The theoretical ramifications, methodological approaches and epistemological goals of this area differ significantly, though there are also striking similarities between some apparently incompatible ideas.

First of all, some scholars have adapted existing representational and theoretical approaches to their work. Connor (2018) uses traditional notation in his transcriptions but laments the lack of temporal precision in its standard duple and triple subdivisions of the beat. Thus, he uses noctuplet notation to achieve a higher rhythmic resolution. As already discussed, Gilbers et al. (2019) use the grid-based representational approach at a high resolution in their analysis of various American rappers. Unlike Connor, however, they deploy their increased resolution specifically to analyse the placement of syllables relative to the ‘normal’ rhythmic categories. That is, they analyse each syllable with regard to whether it is ‘on’ the sixteenth-note position or some amount of time ‘behind’ it, meaning that their increased resolution does not introduce the tiny rhythmic categories of 128th or 256th notes. Rather, beat subdivisions are...

---

38 Interestingly, while Gilbers and colleagues acknowledge that syllables can be placed ‘before or after’ the beat/grid categories (2019, pp. 3, 11), their analyses only indicate the former. The aesthetic ideals of the performers and traditions they analyse might explain this tendency.
used as a scale of measurement instead of, for example, measured milliseconds or samples\textsuperscript{39} in a digital audio workstation.

Ohriner’s work on microrhythm also focusses on comparing the absolute temporal positions of syllables to a grid of subdivisional categories. In his chapter ‘Expressive Timing’ (2018) in \textit{The Oxford Handbook of Critical Concepts in Music Theory} (Rehding & Rings, 2015),\textsuperscript{40} Ohriner uses Kendrick Lamar’s track ‘Momma’ (2015) as one of three examples of his subject, which ‘refers to variation in performed durations among notes represented in a musical score with a single rhythmic value’. The significant difference in expressive timing between the hip-hop track and the other works he analyses (performances of Brahms’s Concerto for Violin and Varèse’s Density 21.5 for Solo Flute) is its rhythmic framework, wherein some aspects of it are ‘mechanically regulated’ (this refers to its programmed/sampled musical background). Thus, Ohriner measures Lamar’s expressive timing using the ways his flow (dis-)aligns with a perfectly quantised timing grid. Ohriner does so via his molecular method, adding a layer above the grid line for circles indicating the ‘continuous onsets’ of the syllables in Lamar’s flow. These circles are connected by lines to the grid below to match them with ‘quantized onsets’. Ohriner expands on the analysis of the same track in ‘Lyric, Rhythm, and Non-alignment in the Second Verse of Kendrick Lamar’s “Momma”’ (2019c) by introducing more levels to the graphic representation of microrhythm.

Alongside his ‘potential quantized notation’ expressed with traditional notation, he plots graphs of ‘syllable duration against onset’, ‘syllable onset within meter” and (a spiral representation of) ‘syllable onset within beat’. Ohriner discusses how different types and degrees of disalignment between continuous and quantised onsets can be indicative of ‘speech-rhythmic’ and ‘music-rhythmic” flows, respectively, and highlights Lamar’s extensive use of speech rhythmicity with a high degree of disalignment. In the same vein, chapter 6 of \textit{Flow} (Ohriner, 2019b) uses these same tools in an analysis of Talib Kweli’s ‘Get By’ (2002) to reveal other techniques for applying expressive timing (or, rather, syllable onsets which are disaligned from quantised onsets). He demonstrates how Kweli’s flows display both consistent repeated disalignments such as ‘phase shift’ (consistently behind or ahead) and ‘swing’ (uneven duration of subdivisions) and non-repeated disalignment.

\textsuperscript{39} Digital audio files are recorded with specific sample rates. A sample rate of 44.1 kHz, for example, means that there is 44100 samples per second. These samples can be used as a scale of measurement in a DAW.

\textsuperscript{40} \textit{The Oxford Handbook of Critical Concepts in Music Theory} is an online publication, with chapters added for several years after its initial conception. This is why Ohriner’s chapter—published in 2018—is part of a publication from 2015.
techniques such as ‘tempo shift’ (slowing down and speeding up again within a phrase) and ‘deceleration’ (a phrase ending with a ritardando, or slowing-down, effect).

**Local form features: larger segments of flow**

While most analyses of rap flow have focused on either the global out-of-time features of whole flows/corpora or the rhythmic features of a fully zoomed-in surface of subdivisional categories and micro-manipulations of these, some analyses have also set their sights on the larger formal units within a verse or flow. Rather than exploring the rhythmic phrasing of cross- and counterrhythms, syncopation and expressive timing, they pay attention to metrical organisation, phrases/lines, hypermetre and the like. While such aspects are naturally part of many analyses of both global features and rhythmic surface structure, those that focus specifically on larger segments of flow showcase interesting representational systems for doing so.

Among the established rap analysts already introduced here, many have conducted some sort of analysis of larger rhythmic segments in time (as opposed to as global features). Ohriner, for instance, analyses how the development of the narrative of the lyrics in Eminem’s ‘8 Mile’ (2002) is amplified by the development of the flow (Ohriner, 2019b, part 6.3.5). Komaniecki analyses how rhyme is used to create formal units like couplets (line pairs) andquatrails (groups of four lines), as well as other rhyme schemes which create larger musical segments (Komaniecki, 2019, section 3.3). Adams introduces a novel approach to the concept of ‘phrase’ in hip-hop (Adams, 2020), defining it as those segments which deviate from the established norms of the genre, specifically in terms of the ways musical metre, syntax and rhyme interact.

None of the above analyses introduces (or needs to introduce) alternative ways of representing the larger musical units around which they revolve. Others do, however, including Katz’s oft-cited conference paper ‘Towards a Generative Theory of Hip-Hop’ (2008). As the title indicates, this paper introduced the use tools from the famous work of Fred Lerdahl and Ray Jackendoff—*A Generative Theory of Tonal Music* (1983)—to analyse hip-hop. Katz employs Lerdahl and Jackendoff’s prolongational and time-span reductions to discern patterns in rap flows’ seemingly limitless capacity to vary rhyme positioning. These reductions are represented as hierarchical trees, common in the field of linguistics (and both Jackendoff and Katz were/are linguists), and in Katz’s work they are used to show, for example, how rhyme connections can cut across rhythmic groupings.
As previously mentioned, Edwards’s (2009, 2013) ‘flow diagram’, a grid system with a low resolution reflecting beats only, is used to present rhyme schemes and phrase/line segments. In peer-reviewed research, John J. Mattessich uses this low-resolution grid to analyse different flows by Kendrick Lamar (Mattessich, 2019). Mattessich very deliberately lowers the resolution from the first track analysed to the second, both because the analytical focus is on bar- and phrase/line-level structural units and ‘to show the flexibility with which Lamar’s flow operates’. I share Mattessich’s focus on the relationship between musical metrical framework and other structuring principles of the vocal track and will revisit his theoretical concept of ‘derivative and generative flow’ in track A4.

More than flow: the breadth of rap analysis

While the focus of this track has primarily been on the analysis of rap flows, a significant body of rap analysis has been aimed at other parts of the musical expression of rap. Several of the publications already mentioned in this track are equally concerned with rap’s musical background, or ‘beat’ (Walser, Krims, Kautny), and another branch of research addresses the flow’s relationship to the beat (Adams 2008, Oliver, Ohriner 2016a). And, alongside the research studying flow—rap’s rhythmic content—is a related subfield which investigates rap’s melodies.

As established in track A1, my definition of flow does not include pitch content or intonation (or whatever else one might call the melodic content of rap). One can rap the same flow with a multitude of different melodies. However, melodic content can be essential to how the rhythm is communicated and/or experienced, and the analysis of rap’s intonation can be essential to illuminating certain aspects of flow. One discipline that naturally combines the analysis of rhythm and intonation is linguistics, and the previously mentioned study by Gilbers et al. (2019) is an example of the linguistic analysis of rap. Tellingly, they include rhythmic information in their melodic analyses, effectively equating ‘speech prosody’ and ‘melody’ in their article (they also include pitch content in their definition of flow). The melodic analysis involves ‘pitch fluctuation’—how frequently there is a significant pitch change, and how large these fluctuations are in semitones. The resulting data are global, however, and do not reveal anything about in-time features like rhythmic patterns or phrasing (nor do the authors want it to). Other linguistic approaches include that of Jan Hognestad (forthcoming 2022) perform analysis using similar tools (creating pitch contours using
analysis software) on single tracks rather than entire corpora. This enables them to look at how intonation is used to demarcate phrase endings, for example.

Music scholars have also adopted these linguistic tools. Ohriner’s article ‘Analysing the Pitch Content of the Rapping Voice’ (2019a) is first and foremost a presentation of a methodological approach—how one can use a combination of automatic data analysis from various software applications and the manual correction of certain parameters to generate useful pitch data for ‘a cappella’ rap tracks (those which are isolated from their musical backgrounds) as well as a graphic representation of important melodic features. Ohriner shows how his approach can tell us something about pitch in rap in general as well as how rap differs from speech and rappers differ from one another. Additionally, he performs ‘three analytical vignettes’ wherein he uses pitch data from single tracks to frame the distinct melodic approaches of three different rappers.

Linguistic analysis software is not the only way to undertake the analysis of melodies in rap. The prolific music blogger Ethan Hein has used Celemony’s Melodyne software (developed for pitch correction/alteration in music production) to visualise and analyse the melodic content in rap tracks (Hein, 2015, 2016). The main punchline of his blog posts is that rap has melody, but there are other interesting observations there as well, including how many rappers mainly stay within a minor third for large sections of rap, and that most rap is (at least to a certain extent) melodically ‘in key’—matching the melody to the musical background.

Analysing rap melodies within a framework of discrete musical pitches is also Komaniecki’s preferred method for presenting different melodic techniques (and categories of techniques) which rappers employ. Graphically, Komaniecki switches between single-line and full five-line staff notation in his transcriptions, using the latter (and ‘normal’ noteheads) where the rap switches into being ‘sung’ while representing rapped sections with cross noteheads on the middle line only. He also adopts a notation with two lines for ‘pitched rhythmic layers’ (2019, pp. 129-132), where the point is to indicate a relative difference in pitch between two distinct layers in a flow rather than specific musical pitches as such. This recalls the method employed by Kautny (2009, 2015), who plots relative pitch levels around a single bar line. A variant of

---

41 Among the different types of software commonly used in linguistic analysis is the freeware Praat, and Gilbers and colleagues, Hognestad and several of the other scholars mentioned in this track use it for at least part of their analyses.

42 Hein is also a published scholar in the field of music pedagogy (Kuhn & Hein, 2021).

43 The switching and/or blurring of the categorical boundaries between speech and rap and song is the topic of track B4 in this thesis, and I will elaborate upon Komaniecki’s work on pitch in rap there.
this can be seen in Rollefson (2017, p. 66) as well, where scratched vocals are notated within a five-line staff but the pitch values are relative.

It seems clear that if one’s goal is to fully dissect a rap track, there are a range of approaches to combine and a vast array of parameters to analyse. Bringing this multifaceted arsenal to bear upon the analysis of one single track, Noriko Manabe illuminates Kendrick Lamar’s ‘Alright’ (2015) from almost every conceivable angle. Her article ‘We Gon’ Be Alright? The Ambiguities of Kendrick Lamar’s Protest Anthem’ (Manabe, 2019) was part of the same symposium on Lamar’s music as the aforementioned articles by Ohriner (2019c) and Mattessich (2019); it celebrated Lamar’s Pulitzer Prize–winning album DAMN.44 In her comprehensive deconstruction of the song, Manabe explores its inherent tension between strength and hope, on the one side, and insecurity, on the other, expressed in the lyrics, the larger context of a concept album, the song’s origins, the different artists’ voices on the song, and the music itself. After analysing the lyrical content, she explores the metrical framework afforded by the backing track and the vocals, which invites a certain ambiguity between the genre-typical 4/4 and a consistent reinforcement of an uneven division into 3/4+5/4 or 3/4+2/4+3/4. The different voices on the track—Lamar’s lead vocals and Pharrell Williams’s ‘hook’ (or chorus), as well as rapper Fabolous’ overdubbing of the backing track—are analysed using both a grid-based system and pitch contour patterns.45 She concludes by unpacking how ‘Alright’ has been used as a protest song at rallies related to the Black Lives Matter movement, and especially the differences in its metrical expression in that context.

As evidenced by analyses like Manabe’s and the sheer range of rap scholarship as a whole across disciplinary and methodological boundaries, there is, as Adams observes, ‘not yet a universal analytical method for hip-hop music’ (2015, p. 121). It may be that this whole notion is unfeasible, given how useful it can be to approach different aspects of rap flow (and other aspects of hip-hop music) from different angles using different tools and theoretical frameworks. My overview of the field centred on rap flow at the expense of the much broader array of subjects in hip-hop studies; while many of the scholars and theories covered in this track will reappear throughout the thesis, I have had to leave others out for now. This overview had another purpose as well, however, which was to illuminate some important

44 The analysed track, ‘Alright’, is not on DAMN. but another of Lamar’s critically acclaimed albums, To Pimp a Butterfly.
45 Manabe uses the aforementioned software Praat for her pitch analyses.
differences in analytical and representational approaches and further the creation of a taxonomy for rap analysis specifically, and the analysis of musical rhythm in general.

Four zoom levels of representation and analysis

When analysts use certain modes of representation to describe, interpret, analyse or sort musical sounds, they are necessarily choosing to focus on some aspects of those sounds.

When one is studying the rhythm of the musical flow of a rapped voice, there are many different aspects of rhythm from which to choose. Different modes of representing the music will highlight different aspects and different features, and the approaches presented thus far in this track can be grouped according to their quarry. In acknowledgment of the fact that the act of re-presenting rhythms on page or screen requires a shift in modality from auditory to visual, I will choose to emphasise this visual dimension when formalising such representational categories. I will call them zoom levels, akin to different lenses on a camera or the adjustment of a microscope.

Throughout this track, we have seen examples of the visual representation of phrase structure (Katz’s time-span reductions), microrhythm (Ohriner’s disalignment analyses), features of the rhythmic surface structure at the level of beat subdivisions and globally (of singular flows or entire corpora) and more. These representational systems can be broadly divided into those which display singular flows as they unfold in time and those which do not, instead superimposing several flows or parts of flows over one another to capture trends and averages. The latter belong to the global zoom level, while the former divide into three zoom levels: the microrhythmic, the quantised and the macrorhythmic.

Following the analogy of the microscope, the microrhythmic zoom level encompasses representations which zoom in on the smallest rhythmic units, such as Ohriner’s and Gilbers and colleagues’ microtiming/disalignment analyses. Representations at this level tend to focus on the type, quality and amount of deviation of the rhythmic events from a temporal ruler, but the most important temporal aspect of microrhythmic representation is that it emphasises rhythmic positions outside categories like beats and subdivisions.

I dub the meso-level of representation the quantised zoom level because the containment of rhythms in strict categories of beats and their subdivisions, rather than the rhythms’ precise nominal time values (as measured in milliseconds, for instance), evokes the quantisation

---

46 As we have previously discussed (and will discuss again in track A6), microrhythm is not simply something which happens on a temporal axis; it is also communicated through sound features.
function used in automated rhythm processing with sequencers, DAWs, notation software and so on. Clarke explains the functionality:

Quantisation is the process whereby the continuously variable event durations of a real performance are rationalised into discrete rhythmic values as represented in standard music notation. For sequencers and notation programs, this stage of processing is required so as to give rise to practical and musically ‘sensible’ representations that might then be read by subsequent performers, or to allow one sequencer track to be coordinated with other tracks. However, the process has also been regarded as analogous to (or even identical with) the perceptual process by which a listener or co-performer parses the rhythmic structure of a performed event sequence, the quantisation process filtering out expressive microvariations so as to reveal the underlying rhythmic structure. (Clarke, 2000)

The quantised zoom level visually represents these categorical ‘discrete rhythmic values’ while (typically) omitting information about variation within these categories (that is, microrhythm). Examples of quantised representation in the literature include traditional music notation and grid-based systems.

Logically, the macrorhythmic zoom level accommodates rhythmic units which are larger than those emphasised at the tighter zoom levels, reducing the amount of visual information to communicate larger rhythmic and formal units more clearly. It encompasses any representational system which highlights phrases, hypermetrical organisation or other large-scale rhythmic features while omitting information about smaller rhythmic categories, like subdivisions of the beat. Katz’s time-span reductions and prolongational analyses are examples of this level (even though they zoom in on individual syllables), but lyric transcriptions are perhaps even better examples, as they indicate some sort of phrase segmentation through their line breaks without complicating this representation with any discrete rhythmic values.\textsuperscript{47} There is no strict boundary between the macrorhythmic and quantised zoom levels in analytical practice, as it is also both possible and common to analyse features which are part of the macrorhythmic level using representational tools from the quantised zoom level. An example of this practice is the use of colour emphasis on transcriptions in traditional notation (as in Komaniecki’s and Kautny’s work, for instance) to indicate phrase and form structures. However, I consider representational practices which

\textsuperscript{47} Lyric transcription and lineation (dividing text into lines) are significant topics in track A4.
specify sub-beat categories to be part of the quantised zoom level, while those where the beat category is the smallest one belong to the macrorhythmic zoom level. Many types of macrorhythmic representation (including lyric transcription) do not even specify the beat category, while some (like most flow diagrams) only approximate beat positions.

Lastly, the *global zoom level* goes out of time, so to speak, to capture the global and ‘grammatical’ features of rap flows. This level emphasises averages, trends and an architectonical overview of a flow (or flows). As far as the zoom lens analogy goes, one could consider the global zoom level to be the widest possible lens, capable of encompassing the entirety of a musical piece (or even a musical corpus). Another analogy might be a brochure wherein the combination of text and illustrations gives a complete overview of an object. Typically, we see this represented visually via graphs containing different types of data, as in the various corpus studies presented earlier in this track.

This zoom-level taxonomy has different uses and implications. For one thing, it might aid in identifying which types of features rap analysts have been most interested in. It might also shed some light on which levels of rap flows are under researched and prompt the exploration of less-studied features and new representational tools to analyse such features. As will be made clear in the following track, it might even be used to link the analysis of rhythmic structure with the human perception of rhythm.
Track A3: Rhythm cognition

Given the definition of rhythm as an interaction between sounding events and virtual reference structures plus the focus on markedness presented in the first track, and the differences in resolution among various representational systems for rap described in the second track, there are three areas of rhythm cognition which crystallise as objects of inquiry here: (1) the interplay between multiple temporal streams, both in sound and in referencing structures/schemata; (2) the categorical perception of rhythm, and the modulation of said categories; and (3) prediction/expectation/anticipation. While the previous track covered how rap flows have been analysed, this track will offer insight into what the analytical tools’ underlying theoretical assumptions are. This is a question of tools and analysis but also of the cognition of rhythm in general terms. The theories and concepts presented in this track also inform the understanding of the rhythmic experience which underpins this thesis.

Streams, structures, schemata

All music involves an interplay among different temporal streams. Even in monophonic music, the single musical voice can be understood as a compound of several different constituents. It contains both melodic and rhythmic information, for example, and while the two are impossible to decouple, they communicate distinctly different structural and expressive information. If there are lyrics present, we must also process semantic, lexical and phonological information. Changes in timbre, volume or other ‘physical’ aspects of a sound are likewise distinct perceptual stream layers, even when they are obviously experienced as constituents of one unified musical or auditory whole. This understanding of musical streams as composite—unified but with clearly identifiable, separable constituents—is at the core of the analysis of rap flows in this thesis.

Albert Bregman (1990) calls these composite streams ‘auditory streams’ and notes that they form as a ‘result of processes of sequential and simultaneous grouping. Sequential grouping connects sense data over time, whereas simultaneous grouping selects, from the data arriving at the same time, those components that are probably parts of the same sound’ (Bregman, 1990b, p. 3). In short, Bregman shows how the temporal organisation and different constituents of sound events inform our perceptual system concerning how to best organise our sounding environment. An important consequence of this theory is that our perceptual grouping of sound must be flexible. In some situations (for example, when conversing with friends in a busy café), we might sort a track of rap music as one single auditory stream.
(perhaps best to be ignored, so that we can focus on our conversation). In other situations, when we can listen to that track closely, we might sort a snare drum, piano and voice(s) into different streams, *attending* more closely to a single one or the specific interaction between two of them as we parse the musical information.

Returning to an understanding of rhythm as an interaction between Danielsen’s concepts of *figure* and *gesture* (see track A1), this coming-together of sounding musical utterance (the gesture) and virtual reference structure (the figure) can shed further light upon the composite nature of Bregman’s auditory streams. Just as a figure is a virtual representation of the sounding gesture, Bregman’s auditory streams are specific mental representations of sensory data. One gesture can be represented by many different figures, and, as a result, many different features of the actual sounding gesture can be emphasised. Or, alternatively, the gesture can be looked at from different angles, each one illuminating slightly different aspects. Continuing the visual metaphor, one might think of these different figures as lenses emphasising different combinations of layers in the composite auditory stream. Our mental representations might be holistic in the sense that they turn a mass of sensory data into a single unified auditory stream, but they are also flexible in that, depending on the context, they emphasise different features of the whole.

Where Danielsen links figures to *reference structures* or *schemes*, both Bregman and Huron (2006, chapter 12) write about *schemas* (alternatively conjugated ‘schemata’). Where Danielsen allows the concept of reference structures to encompass both style- or genre-specific patterns as well as universal encultured features such as metre and emergent patterns from the specific musical stream, Huron delimits his use of schemas (and ‘schematic expectations’). While still containing ‘style’ or ‘genre’ as well as other ‘broadly enculturated patterns of events’, he defines emergent patterns within the musical context as ‘veridical expectations’ and ‘dynamic expectations’ for long-term and short-term patterns, respectively (Huron, 2006, p. 231). I have little need for an active use of this taxonomy and will use Danielsen’s ‘reference structure’ throughout this thesis.

**Categorical rhythm perception**

Kvifte (2004) makes an important distinction between the analogue and digital components of central musical parameters.48 A pitch class is categorically different from another pitch class (a Bb is never a G), and Kvifte refers to this as the digital aspect of pitch. Yet each particular

---

48 Kvifte also references another common terminological pair for the same concept: continuous and discrete.
instance of a pitch class might have very different analogue properties which affect our perception of pitch (such as wide vibrato, a slightly out-of-tune instrument, effects such as chorus and so on). Such variation does not necessarily influence our perception of pitch class, though. The G is a G until it becomes a Gb/F# or a G#/Ab. Rhythmic categories are similar. We often write, print or talk about different note values as though they were precise temporal entities. A quarter note is a quarter note, which has a specific temporal position, and each quarter note is subdivided into two eighth notes, which are thus, in one sense, ‘equal’ (they belong to the same category at some level, even though there might very well be a categorical differentiation between on-beat and off-beat eighth notes at another level). These are digital aspects of the rhythmic categories of note values, or as Bengtsson (1987) calls them, ‘time-value classes’. There are, however, many examples wherein ‘time-value classes’ are clearly uneven, therefore displaying different analogue qualities. Swing, ‘pushed’ snare hits in some styles of R&B, any kind of rubato performance, and uneven meters in Scandinavian fiddle music are just some examples wherein note values display great flexibility. Two eighth notes do not have to be equal (in some styles of swing, one eighth note can even be twice as long as the next), but they are still experienced as being part of the same time-value class. They are both eighth notes, but they are unequal. Of course, categories may be nested within one another depending on the context in which they are experienced/interpreted as well. A half note and a dotted half note may both belong to the category ‘long’, and the difference between them might be analogue in the categorical context of long/short but digital in the context of time-value class.

The fact that note values are categories rather than absolute temporal durations does not make them mere approximations. Their function is precisely to work as categories in our perception and/or understanding of sounding music. When we perceive an auditory stream, our mental representations, or figures, might emphasise the digital aspects of the rhythm—in some cases, that is, the significance resides not in whether the eighth notes are even or uneven but in whether they are eighth notes or something else (in other words, whether they belong to the category ‘eighth note’). Such a categorical differentiation, like that between ‘part of the beat’ and ‘not part of the beat’, is core to some of the theoretical concepts which inform the

---

49 And ‘note class’ is simply one type of pitch category. The same sample pitches can belong to other categories within different structuring frameworks. In the context of a major scale, for example, the note E is categorically ‘a third’. More specifically, it is a major third when the resolution of the categorical framework is chromatic rather than diatonic. In the context of the chord progression of Am7–D7–Gmaj7, the notes C, F# and B will (in succession, not simultaneously) all belong to the category of ‘third’ (but the C will not be ‘major third’).
analyses on side B of this thesis, several of which will be fleshed out in track A6. Whether something belongs to a specific auditory stream or not (Bregman’s *stream segregation*) is also a categorical question. Other times when we perceive an auditory stream—for example, in some of the microrhythmic analyses presented in track A2 and performed on side B, our figures might emphasise the analogue aspects of the rhythm instead.

In his article ‘Levels of Structure in the Organization of Musical Time’, Eric Clarke (1987) introduces three distinct structural levels in the organisation of musical time: the low-level category of ‘expression’, the middle-level category of ‘rhythm and meter’ and the high-level category of form’. The first level is, according to Clarke, ‘perceptually represented as departures from canonical proportional values (…) and experienced as expressive rather than durational effects’ (Clarke, 1987, p. 233). If we relate this level to Kvifte’s digital/analogue distinction, we could say that the expression level is about the perception of analogue qualities *within* time-value classes. It is not about which note value a rhythmic event is but rather about the variations within the perceptual boundaries of a note value (or, more precisely, a time-value class).[^59] I discussed the relationship between rhythmic expression and rhythmic structure in track A1, and that relationship is exactly what Clarke describes via his categorisation.[^51] The category ‘rhythm and meter’ is ‘perceptually represented as a collection of grouped durational equalities and inequalities organized around a metrical framework’ (Clarke, 1987, p. 233). This is the perceptual level of digital time-value classes. Equivalent, but not necessarily equal, rhythmic events are categorised, and the groupings of these are perceptually sorted. The high-level category of ‘form’ is ‘perceptually represented as a structure of hierarchical relations, constructed by means of memory processes and perception, and distinguished from level 2 [rhythm and meter] structures by exceeding the length of the perceptual present’ (Clarke, 1987, p. 233). The structural characteristics and properties of ‘form’ are not significantly different from those found in ‘rhythm and meter’ (metre, groupings, and so on), but, as Clarke notes, ‘a form is, by virtue of its overall duration, a structure that is constructed out of perceptual and memory information’. ‘Form’ is shaped not only in the immediate unconscious perception but also in the ways in which this perception

[^59]: Of course, if the categorical distinction is at the level of ‘long/short’, for example, the question of *note value* can be an analogue variation. It is specifically within a context wherein the note values are the categorical units which are experienced/analysed (which is the case for most of the analyses referenced and performed in this thesis) that it is strictly true that ‘it is not about which note value a rhythmic event is’; more generally, it is more accurate to say that ‘it is not about which categorical time-value class a rhythmic event belongs to’. However, in my applications of Clarke’s concept, I find this distinction to be complicating rather than enlightening.

[^51]: Clarke expands upon this relationship in his ESCOM conference paper ‘Categorical Rhythm Perception and Event Perception’ (Clarke, 2000).
interacts with our memory of the preceding perceptual segment (which also informs our predictions about what is likely to come next). Both Clarke and other scholars write primarily about ‘goal-oriented’ classical music and focus on cadences and other structurally significant features of harmony in their exploration of musical form. In the analysis of rap flows and other cyclic, groove-based musical expressions, of course, harmonic features are much less prevalent or prominent. Yet even in rap, we still find structural blocks larger than the meter, and in my view this is one of the more interesting underresearched parameters of rap flows.

**Categorical rhythm perception and the zoom levels**

The aforementioned zoom levels—the different categories of representation in rap analysis introduced in the previous track—overlap nicely with Clarke’s structural levels of rhythmic perception. ‘Form’, ‘rhythm and meter’ and ‘expression’ are analogous to the macrorhythmic, quantised and microrhythmic zoom levels, respectively. This overlap has implications for how we should understand analyses performed at the different zoom levels, as well as for how we should conduct our analyses and (even more importantly) how we can create new representational systems which effectively emphasise the rhythmic features in which we are interested.

As previously stated, representation at the *microrhythmic zoom level* tends to focus on showing the type, quality and amount of a rhythmic event’s deviation from a temporal ruler. This temporal ruler is often (but not always) a grid of subdivisions of isochronous beats. If our temporal ruler is an adequate representation of a ‘true’ perceptual categorisation, perceptual and representational features will correlate very well, as both are defined by their deviation from strictly defined note-value positions (or time-value classes). Thus, if we adopt Kvifte’s digital/analogue distinction, the microrhythmic zoom level is a visual representation of analogue features of time-value classes. It is not quite so straightforward, however, as there might be blurring of categorical borders, overlap between different categories depending on context and/or temporal ordering, and even competing or interacting reference structures.

---

52 For a more in-depth exploration of the interaction between perception and memory, as well as the term ‘perceptual present’, see Clarke (1987) and Michon (1978).
53 Notably Lerdahl and Jackendoff (1983).
54 As I will discuss further in track A6, temporal rulers can (and should) also be derived from analyses of the work or genre being analysed. Good examples include analyses of music from traditions displaying unisochronous metres, such as certain Scandinavian fiddle-music traditions—see, for example, the work of Kvifte (2004, 2007) and Johansson (2010a). Genres like the Vienna waltz and samba also display unisochronous metres requiring specific temporal rulers which are themselves flexible—see, for example, Haugen and Danielsen (2020).
which produce categorical overlaps and/or disagreements. The interesting challenges which arise in the analysis and representation of microrhythm—and particularly those pertaining to ambiguity and the listener’s emergent interpretation and reinterpretation of a rhythmic structure in passages with significant (analogue) expressiveness—will be a topic of discussion in track B3.

The quantised zoom level encompasses the digital counterpart to the analogue qualities captured by the microrhythmic zoom level. Visual representation can be used to categorise rhythms as patterns of specific note values, as is the case for traditional notation and the various grid-based systems. Note that, when using or assessing this level of representation, it still allows room for plasticity as to what those note values actually are when transformed into sound (that is, the mental and graphic figure is quantised, even though the acoustic signal might not be). Not all eighth notes have the same duration, but in most cases they are still eighth notes. Often it is the digital (that is, the categorical) time-value class of a rhythmic event which is most relevant to the analysis at hand, and the need for structural clarity might necessitate the omission of information about most of the analogue aspects. Quantised representation is a suitable tool for the task, then, with the caveat that it does not imply quantised sound.55

The macrorhythmic zoom level, again, is the representational equivalent of Clarke’s perceptual category of ‘form’. Whereas the microrhythmic zoom level shows analogue features of time-value classes and the quantised zoom level shows sound sorted according to such (digital) classes, the macrorhythmic zoom level shows large-scale structural features which do not necessarily need to specify discrete time-value classes (like note values) or which at most use only larger and ‘broader’ classes (such as approximate beat positions or bar numbers). Rather than capturing the precise execution of rhythmic figures, representational systems at the macrorhythmic zoom level concern themselves with longer structural segments or connections—those which are perceived through the aggregation of information from the perceptual present and beyond it (memory and anticipation). While typical ‘form’ features such as hypermetric structure or the positioning and segmentation of phrases are also visible at the quantised zoom level, there are still several reasons to ‘zoom further out’. If one wants to highlight a specific structural feature, the best choice of visual representation should centre

55 This is one area where the analysis of recorded music and the analysis of score-based music can be thought of as similar though also opposite. In score-based music, the prescriptive score is quantised but the expectation and common practice are that performances based on the score will be expressively de-quantised.
upon it rather than concern itself with any additional features. Just as we omit representation of many analogue aspects of rhythmic events at the quantised zoom level, we omit any representation of most time-value classes at the macrorhythmic zoom level. I will explain my own sense of these structural levels and apply them to sample analyses in track A6.

In relation to Clarke’s perceptual categories, the last of my zoom level categories—the global zoom level—concerns itself less with temporally specific perception and more with those rhythmic features which can be observed outside of the ongoing temporal ordering. In this thesis, I am interested in what happens in composite auditory streams when they actually take place—that is, as they are experienced—not in an aggregated overview of a larger corpus; therefore, my analyses will not rely upon any global zoom level representation. However, this zoom level does supply invaluable insight into both the genre-specific reference structures and the overall development of rap, and I will draw extensively upon others’ research at this zoom level throughout my discussions.

**Prediction, attention, markedness and expected convergence**

Going back to the titular mega-question, we will now move from the ‘what’ into aspects of the ‘dope’. While the aspects of prediction and attention are crucial to rhythm perception, they are also crucial to markedness—something standing out, demanding our cognitive processing power—which is itself crucial to making shit dope. Auditory streams are temporal, sequential phenomena, and we experience them as constantly changing, continuously developing perceptual virtual objects. Through both pre-existing and dynamically emergent reference structures, we develop our predictions about what will happen (or our general expectation or anticipation that something will happen) in the auditory stream to which we are listening: ‘To predict is to anticipate the future based on the past. This requires the existence of an internal model, which captures learned associations and regularities upon which predictions can be formed’ (Schröger, Kotz, & SanMiguel, 2015, p. 4). Likewise, our reactions, evaluations and appraisals of what has already happened in said stream will inform and alter both the reference structures we apply and our sense of the very nature of the auditory stream itself. Perhaps it is not a single composite stream but several separate ones, for example. No matter what, our re-evaluation of the nature of the stream(s) does not change the fact that, up to the point when we change our minds, our abiding notions of the stream and its internal, emergent rule sets may have been incompatible with our newfound perceptual knowledge. Our experience of musical rhythm, in other words, is incredibly dynamic, and in our engine room sits a prodigiously capable ‘prediction machine’ (Clark, 2013, p. 181). It is also worth noting
that, in cognitive psychology research, music unfolding in time represents an especially
fruitful type of experimental stimulus as ‘the implicit structure of music provides a particular
situation in which very strong auditory expectations arise’ (Schröger, Marzecová, &
SanMiguel, 2015, p. 648).

Attention, then, is very important to our theoretical framework, since the interaction of many
types of expectation, many different reference structures and the many layers of the composite
auditory stream demands that many different things go on simultaneously in our brains.
Obviously, not all of these things can receive the same amount of ‘processing power’ at the
same time: ‘A classic distinction of how auditory information can be selected is whether the
selection happens voluntarily (endogenously, top-down driven) by the intention of the
listener or whether it happens involuntarily (exogenously, bottom-up driven) without the
explicit intention of the subject’ (Schröger, Marzecová, et al., 2015, p. 651). This distinction
helps to explain how and why we attend to different layers (or combinations of layers) in a
composite auditory stream, and why this attention suddenly switches. For example, if we are
attending specifically (voluntarily, ‘top-down’) to the rhythmic pattern of the syllables in a
flow, we might be ignoring the semantic content of those syllables. But if something in that
semantic content is sufficiently marked, therefore disrupting (or causing a ‘prediction error’
in) our ongoing subconscious processing of it, our attention might shift between rhythmic and
semantic layers through an involuntary, ‘bottom-up’ response—attention and prediction, that
is, are inextricably linked. Of course, this entire model is a very simplified take on what goes
on in our brains when listening to music. For one thing, different sensory data are processed
in different parts of the brain, and (as Huron stresses throughout ‘Sweet Anticipation’ while
linking expectation to emotional effects) the same data are also processed by different systems
and via different pathways with different requirements for cognitive processing. In addition,
the brain does not neatly divide a complex auditory signal into distinct layers. However, there
is physiological evidence of the impact of prediction errors (that is, markedness) on both
attended and unattended listening and the ways in which our selective attention amplifies our
various responses to auditory signals (Schröger, Marzecová, et al., 2015), and one can
theorise about how this might affect which rhythmic features we will experience as aesthetic.
The nuances of complex musical contexts are very hard to model in electrophysiological
experiments, so there is a place for simplified explanatory models. As will hopefully be clear
in the coming tracks, the model of the composite auditory stream and its many interacting
layers has been very useful for uncovering and interpreting *dope shit* in my analyses, and I believe it can be used in a similar fashion by others and in other contexts.

In a theoretical framework wherein top-down attention and following (and predicting) the content of layers of the composite auditory stream interact with bottom-up attention via prediction errors and markedness, rhythmic techniques can be explained via the concept of *expected convergence*. This is the insight that many rhythmic effects and techniques are based on the interaction between layers of the composite auditory stream and, by extension, the expected convergence of features within those layers. As these layers are both separable and inextricably tied together, it is a given that many of them represent reference structures (both pre-existing and emerging as the music unfolds) which are created and/or informed by multiple other layers in combination. For example, reference structures like musical metre and poetic lineation (both of which will be discussed at length in the following track) are layers which are continuously informed by what happens in other layers (the vocal and drum parts, for example). In this thesis, concepts like *line/bar coincidence* (tracks A4, A6, B1) and *metrical anchors* (track B3) are rhythmic structuring features which result from expected convergence (between musical metre and lines, and stressed syllables and metrically strong beat positions, respectively).

Most times, the impact of expected convergence upon our listening experience is either obvious enough to be unspectacular or simply useful for strengthening our sense of the musical structure as it unfolds. Other times, however, there is such a clear break in our expectation that an event becomes clearly marked and ‘sticks out’. Many of the concepts related to expected convergence which will be explored in this thesis, in turn, are rooted in situations where contradictory information from various layers of the composite auditory stream creates rhythmic effects which are either ambiguous or hint at some sort of ambiguity. This rhythmic ambiguity need not be experienced and interpreted in a conscious manner, and exactly *what happens* rhythmically might not be evident from a single listening experience. Oftentimes, analyses of features related to expected convergence originate in an initial sensation of *I know something happened there, but what on earth was it?*; the subsequent interpretation of what caused that feeling will be the result of repeated re-listening and even the examination of out-of-time close-ups of waveforms and spectrograms. Still, my hypothesis is that many rhythmic techniques employed in rap flows are based on a play with expected convergence, and whether the artist does so consciously or subconsciously—and
whether the listener experiences it consciously or subconsciously—is of secondary importance.

Summary

In this track, I have presented some existing theories of rhythm cognition which inform my approach to rap analysis. First, from the concepts of auditory streams/stream segregation (Bregman) and of musical reference structures/schemata (Danielsen, Bregman, Huron), I distil my own concept of composite auditory streams to capture the fact that we perceive musical sound as a whole consisting of multiple interacting layers, as well as my concept of categorical rhythm perception and the way in which we hear both discrete rhythmic categories and (continuous) variations within those categories. I also adopt Kvifte’s terminology of digital for the discrete categorical and analogue for the continuous or ‘expressive’. One important takeaway is that the different categories of categorical rhythm perception in Clarke’s ‘Levels of Structure in the Organization of Musical Time’ (1987) correspond to the different categories of representation in rap analysis—that is, the zoom levels I introduced in track A2. This correspondence provides a theoretical foundation for what representation within these categories can and should highlight, setting the stage for the presentation of my own representational methods in track A6.

Following this, I also introduced the concepts of prediction and attention and related them to the recurring idea of markedness as an indicator of something being dope. Prediction (and attention) is also central to the final concept of expected convergence—the idea that many rhythmic effects and techniques are based on the interaction between layers of the composite auditory stream and the expected convergence of features within those layers. Among the reference structures which constitute layers of the composite auditory stream of a rap flow are musical metre and poetic lines, and their nature and interaction are the topics of the next track.
Track A4: *Metre on metre or verses in verses?*

Descriptions of rap will often emphasise that it is both music and poetry. Of course, it can be argued that poetry is a type of musical form in itself, but it should be evident that rap operates with a rhythmic language which is quite different from most other types of poetry. The ways in which these two artistic expressions are typically engaged also differ in terms of their modality. Rap as an artistic product is sold and consumed as audio, whereas poetry is most often found in book form (recordings are by no means uncommon, but they are not the norm). The analysis of music and the analysis of poetry is equally similar but separate. The fields of music and literary theory both engage in rhythmic analysis and employ much of the same terminology. However, their respective applications of this terminology from different angles have some nuanced theoretical implications for the analysis of rap in particular. For example, the concepts of *metre* and *verse* are key to deciphering rap’s structure(s) from both a musical and a literary angle, as are the intersections of the respective disciplines’ applications of these terms. This track explores the meeting of musical and poetic metre—their similarities, differences and interplay—and the ways in which it is central to the structure and aesthetic expression of rap.

**Metre: Musical and poetic, experiential and analytical**

Of all the virtual reference structures we use to organise our experience of compound auditory streams, *metre* is among the most salient. It is also an excellent example of a reference structure that is not explicit in the acoustic signal but instead virtually (pun intended) universally acknowledged as a part of a musical (and, as we will see, poetic) cognitive framework. As a cognitive phenomenon, however, it is remarkably difficult to pin down in a simple definition. It does not help the situation that there are both striking similarities and irreconcilable differences between the term’s applications in music and in poetry. To begin to make sense of them, I will turn to Mats Johansson’s exploration of the musical applications of the term in his aptly titled 2010 article ‘What Is Musical Meter?’.

Four different, more or less overlapping, perspectives on meter are discussed: (1) Meter as a measuring device, specifying the temporal relationships between rhythmic units and levels (beats per measure, etc.). (2) Meter as an imposed or inferred accentuation pattern (strong-weak-weak, etc.). (3) Meter as an emerging property of the listener’s engagement with the unfolding music, implying that there is no pre-existing neutral grid in relation to which musical sounds are rhythmically structured.
(4) A formulaic conception of meter: a stylistically coded (i.e., culture-specific) notion of sameness resulting from a continuously ongoing process of trying out different, but metrically equivalent, rhythmic designs. (Johansson, 2010b, p. 41)

There is no obvious coming-together of these approaches under a single unified theory of metre, musical or otherwise, and Johansson notes that some of them might even be mutually exclusive. He also problematizes his own formulaic conception of metre by saying that it might not really describe ‘metre’—‘what is referred to would more correctly be termed groove or style’. He concludes by articulating a succinct and precisely formulated common denominator to the different approaches to musical metre: ‘meter may be seen as a frame of reference within which musical events are made sense of’ (Johansson, 2010b, p. 56)

Metre as a frame of reference is also central to Danielsen’s definition of the term, the discussion of which is summed up under the subhead ‘What Is Metre?’ in her 2018 article ‘Pulse as Dynamic Attending: Analysing Beat Bin Metre in Neo Soul Grooves’ (Danielsen, 2018). If we start at the end of Danielsen’s discussion, she defines musical metre as ‘the virtual schemes that correspond to the relatively regularly recurring pulsations at different frequencies (tempi) in the listener’ (Danielsen, 2018, p. 181). This definition complies with the more traditional musicological understanding of metre. In Justin London’s Hearing in Time (2012), he lays out a comprehensive theory of musical metre based on traditional music theory and insights from music cognition. While metre is not ‘in’ the sounding rhythm, it is an emergent property of it, created in tandem with the experience the listener already possesses. Or, as London puts it: ‘We fit, so to speak, patterns of events in the world to patterns of time we have in our minds (and, as we will see, our bodies)’ (London, 2012, p. 4).56 London also emphasises the fact that metre necessitates more than one level of periodicity. Perceiving, attending and entraining to a series of pulsations is the work involved in beat perception, and when we organise these beats into larger periodical units or groups, we have metre.57 In fact, London notes that we tend to prefer three or more levels of periodicity ‘as this provides an attending framework that allow the listener to track rapid, moderate, and relatively slow event onsets, and these correspond to subdivisions of the tactus, the tactus level itself, and a higher-level ordering of beats into measures’ (London, 2012, p. 16).

56 The embodied nature of rhythmic cognition—in the sense that our entire bodies, not merely our brains, are central to our ability to perceive rhythm—has been well established in the field of cognitive psychology. See Shapiro (2011) for an overview.

57 A summary of the research on the differences in beat and metre perception processes in both humans and animals can be found in Fitch (2013).
This way of defining experiential metre is by no means restricted to music scholars, and Richard Cureton’s definition echoes Danielsen and London perfectly while also accommodating modalities other than the auditory: ‘Meter represents our rhythmic response to (relatively) regular pulsations in a perceptual medium’ (Cureton, 1992, p. 123). ‘Our rhythmic response’ is directly analogous to a virtual reference structure, which, as we have established, is no less real than the acoustic signal to which it responds. In terms of Danielsen and London’s shared notion that reference structures precede the sensory signal, for example, it appears in Cureton’s presentation of traditional foot-substitution prosody as well, where there is a dichotomy between ‘actual prose rhythm’, which is the specific structure of the lexical stresses in the poem, and the ‘silent metrical rhythm’ one imparts to the poem (Cureton, 1992, p. 7). Another related conceptualisation of our cognitive organisation of rhythm from the field of poetry is Roman Jakobson’s distinction between ‘verse design’ and ‘verse instance’ and ‘delivery instance’ (Jakobson, 1960, pp. 364-365). These could well be thought of as virtual reference structures and their actual counterparts.

An important aspect of all these theories is that they not only describe a process of evaluation whereby we organise what we have heard but also indicate that we use this continuous evaluation and entrainment—as well as our pre-existing reference structures—to infer what will happen next. Cristopher Hasty calls this ‘meter as projection’, or ‘the process in which a mensurally determinate duration provides a definite durational potential for the beginning of an immediately successive event’ (Hasty, 1997, p. 84). Metre is both in the past (reminiscent of Clarke’s category of ‘form’ discussed in the previous track), the present (which Hasty emphasises in describing ‘meter as process’) and the future.

Thus far, my focus has been on metre as an experiential phenomenon, but the way in which we as scholars, listeners or readers engage with metrical art tends to be through some sort of more or less formalised systems or discourses involving the metrical. Danielsen makes it a point to distinguish these analytical systems of metre from the strictly cognitive processes of our physical apparatus, with the caveat that the analytical and experiential modes of metre are not necessarily readily separable. While the experiential mode of metre involves, as we have established, the reference structure or structures we apply to our metrical experiences, the analytical mode is ‘the standard with which we measure and map rhythmic events in

58 ‘Verse instances’ can vary against the template (or reference structure) of ‘verse design’ on a line-to-line basis. Jakobson distinguishes ‘delivery instance’—the performed variation—from ‘verse instance’.
analytical and compositional representations of music’ (Danielsen, 2018, p. 180). This work need not be limited to music, of course, or to poetry.

Looking at both music theory and poetry, the most common analytical frameworks and discourses in use echo the properties of metre which have been established here. Specifically, we find multiple formalised levels of periodicity in both traditional music notation and foot-substitution prosody and its related systems of metrics. There are clear parallels with regard to both the tactus-level events and the groupings of these events. In musical metre, beats are grouped into a larger metrical unit, called a bar, while in metrical poetry, one counts the number of syllables or stresses within a line. These beat-level events are also amenable to multiple different means of subdivision. To distinguish between metrical units at the beat level and groupings of beat-level units, I use the term *span* for the latter—that is, for musical bars and poetic lines, as illustrated in figure A4.1.

<table>
<thead>
<tr>
<th>Subdivisions of beats</th>
<th>Musical metre:</th>
<th>Poetic metre:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subdivisions</td>
<td>Syllables of feet</td>
</tr>
<tr>
<td>Beat-level units</td>
<td>Beats</td>
<td>Feet</td>
</tr>
<tr>
<td>Groups of units: spans</td>
<td>Bars</td>
<td>Lines</td>
</tr>
</tbody>
</table>

**Figure A4.1:** Correspondence of musical and poetic metrical-structural categories.

While these categories correspond nicely and have similar functions in the metrical organisation of music and poetry, we will see in this track that they do not necessarily coincide when musical and poetic frameworks come together in, for example, a rap flow.

**Metre on metre? Coinciding units or coinciding spans?**

One scholar exploring the idea of an overlap between poetic and musical metre within the context of rap is Adam Bradley, who goes so far as to state that ‘the beat [that is, the musical background] in rap is poetic meter rendered audible’ (Bradley, 2009, p. xvi). Just as rap is a meeting of musical and poetic expression, metre is a meeting of musical and poetic structure. However, where Bradley’s statement clearly implies that poetic and musical metres come together and overlap snugly, I would argue that the relationship is much more complicated. Rather than a single joint or shared metre, that is, there is an interaction between several

---

59 The similarities among the rhythmic groupings in language, poetic metre and music and the cognitive computations of these rhythmic structures are explored in Halle and Fabb (2011).
different aspects of the related phenomena of musical and poetic metrical structures. The question of which parts (or levels) of the metrical frameworks converge, and whether these can even be considered ‘metrical’, is not straightforward. In addition, there is the question of how literally Bradley intends ‘the beat’—the musical metre as expressed/inferred by a musical background—to be considered equal to the poetic metre; his discussion of the topic is much more nuanced than his initial statement suggests. Nevertheless, I will use it to kick off the following discussion of the relationship between musical and poetic structure and analytical frameworks.

In the strictest sense, the statement ‘the beat in rap is poetic meter rendered audible’ could be taken to imply that the metrical units at the pulse/beat level—the musical beats and poetic feet—coincide. While this might occasionally be the case, it is patently not a signature characteristic of rap. This can be demonstrated with simple negative evidence, as in figure A4.2.

![Figure A4.2](image)

**Figure A4.2**: Transcription of bars 1–4 of J Cole’s verse on 21 Savage’s ‘a lot’ (2018). From 2:37.

In bars 1, 2 and 4, the stressed syllables and rests coincide with the musical beats, as one would expect if poetic and musical metres were ‘the same’ or perfectly overlapping. However, bar 3 diverges completely from this one-to-one relationship by organising its syllables in a duple grouping over a triple-subdivided musical metre. This results in six stressed syllables over four musical beats, and the cross-rhythm even displaces stressed syllables from the second and fourth beats. 60 This means that the pulse-level metrical *units*—the musical beats and poetic feet—do not form a coherent, converging structural layer in rap; instead, they are two different but interacting structural units from equivalent levels of two different metrical frameworks.

---

60 Note that there are important nuances to the segmentation of both musical and poetic metrical spans (bars and lines) in this example, in that either or both could be heard in ‘double-time’ compared to the metrical framework(s). This aspect of rap’s (and specifically this example’s) metrical possibilities is fleshed out in track A6.
Another interpretation of Bradley’s statement might be that the metrical equivalence is not at the pulse-level metrical units but rather at the larger grouping level of the metrical spans—musical bars and poetic lines. This would mean that a bar is equal to a line in terms of duration but not necessarily rhythmic content. This may, in fact, be Bradley’s stance, as he relatedly calls for transcribing ‘rap verses in such a way that they represent on the page as closely as possible what we hear with our ears’ (Bradley, 2009, p. xviii)—that is, by segmenting the lines of lyrics according to the musical metre. Right away, however, we face certain interesting questions about the most fundamental visual representation of rap vocals: transcribed lyrics.

Some scholars argue that rap verses are in their very nature texts which require transcription in order for the listener to be able to experience all their nuances a verse communicates, or that the lyrics are only completely decipherable when they are transcribed. Pate (2010) calls rappers ‘rap/poets’ and their raps ‘rap/poems’ and advocates analysis which is similar to that conducted on written poetry. Likewise, Caplan (2014) explains how rap flows stand in contrast to the idea that written poems should be read aloud to be fully experienced:

A certain kind of hip-hop lyric, though, demands nearly the opposite progression. Crafted according to the demands of musical performance, it seeks to be transcribed and considered as a silent written text. Responding to this call, students of hip-hop construct numerous website databases that catalogue thousands of lyrics, so viewers can study them as words set into stanzas. ‘Do you fools listen to music or do you just skim through it?’ asks Jay-Z. To ‘listen to music’ intelligently, then, is to write it down, to read it attentively. Encouraging this process, many hip-hop artists welcome archival efforts to document and analyze their work, recognizing that this study validates their artistry. (Caplan, 2014, pp. 15-16)

These archival efforts by fans are often the only access anyone has to any written version of rap lyrics, as it is not particularly common for rappers to publish their lyrics anywhere. Websites like genius.com, where the archival work is crowdsourced to literally millions of users, take on an added level of authority, as their transcriptions and annotations are used by the streaming software Spotify as a visual accompaniment to the music.

---

61 It is not generally thought that one must catch every reference, recognize all the symbolism or even just understand all the words in a rap track in order to enjoy it. However, rappers do like to include esoteric references and saturate their verses with double and triple entendres, and it typically requires both multiple listenings and/or some amount of reflection or discussion to pick up on many of the lyrical details.
These transcriptions are a form of visual representation that presents suggestions for only certain structural dimensions such as segmentation into lines and larger formal units. These are the ‘verses in verses’ to which my track’s title alludes. In literary theory, ‘verse’ typically refers to the structural unit of a line, while music researchers (and literary scholars working with songs) refer to this unit as a ‘stanza’—or, in the words of Attridge: ‘A verse can mean either “a line” or “a stanza”’ (Attridge, 1995, p. 5). While Attridge avoids this potential terminological confusion by not using the term at all, I follow colloquial rap discourse and use the term verse for the larger formal unit, whereas I follow Attridge in the use of ‘line’ for poetic metrical spans. Any transcription is in its very nature an interpretation of what the lines in a rap verse actually are. In transforming rap lyrics into written poetry, one is necessarily dividing a stream of sounding lyrics into poetic lines, but this practice has no especially clear dictates or philosophy underpinning it, and one might well wonder whether a given transcription actually remains true to the transcriber’s experience of the performed lines or accommodates other concerns, such as presentability on the page or screen, evenness of line length and so on. Adherence to syntax over aspects such as the actual positioning of the rhymes in the rap might also impact the transcription, because punctuation and capital letters likely correspond to line endings and beginnings but perhaps not to the delivery. Most importantly, we usually do not know what the transcriber was thinking while choosing how to segment the text. On lyric-aggregator websites such as the aforementioned (and popular) genius.com, it is typical for multiple transcribers to contribute to a single verse (or even line!), and there are no guidelines on the site as to how one should lineate (though there are many, many guidelines concerning other aspects of the page’s functionality, and particularly the annotation of interpretations of the lyrical content).

Bradley’s commitment to using musical metre as a guideline for lineation represents an attempt to formalise rap transcription with two admirable goals in mind. First, consistency between transcriptions (and transcribers) is clearly helpful to comparative analysis. Second (and Bradley’s main argument for the practice), it reveals the musical metrical structure, which is otherwise lacking in a pure text transcription. As will be made clear throughout this track, I believe this commitment is nevertheless inherently flawed, but its origins are valid.

---

62 This means that my track’s title could be called ‘lines in stanzas’, but (1) this does not sound nearly as good, and (2) it fails to represent the ambiguity and floating boundaries between different structural units which is the running theme of this dissertation.
The convention of line/bar coincidence: Should it dictate rap transcriptions?

One of the global features of rap which stands out from the various corpus studies discussed in track A2 is line/bar coincidence. As hip-hop music developed, rappers came to organise the lyrics via the positioning of simple end rhymes toward the end of each bar of music, either on or in close proximity to the fourth beat of the ubiquitous 4/4 musical metre. This tendency is clearly visible in Condit-Schultz (2016c), for example, where a corpus-wide distribution of stressed rhymed syllables clearly congregates on or anticipates the fourth beat of the bar (as discussed in track A2). Ohriner (2016a) duplicates these findings but through a slightly different methodological approach, looking at ‘line endings’ and ‘phrase endings’ rather than strictly stressed rhymed syllables. In advance of the closer look we will take at the concept of ‘line’ later in this track, it is worth noting that Ohriner has a very particular definition of both ‘phrase’ and ‘line’ within the scope of that specific article: ‘All the syllables said in one breath comprise a phrase of rapping. A line of rapping is a segmentation such that there are as many segments as there are measures in the verse and each segment is relatively closed syntactically’ (Ohriner, 2016a, p. 158). Here, the line/bar coincidence, or at least a one-to-one ratio of lines and bars, if not a perfect overlap, is a prerequisite for the definition of line.

Given the rap convention of a structural convergence between metrical spans, it seems sensible to transcribe rap lyrics with line breaks which adhere to musical metrical boundaries, especially if one’s goal is a consistent practice which skirts, as much as possible, the analyst’s subjective choices. In practice, however, there are several complicating factors. First of all, how strictly should one adhere to the musical metrical boundaries? The following passage from Nicki Minaj’s ‘Anaconda’ (2014, from 0:07) appears straightforward but eludes transcription with a strict adherence to musical bars:

Boy
toy named Troy used to live in Detroit. Big
dope dealer money, he was getting some coins. Was in
shootouts with the law, but he live in a palace. Bought me
Alexander McQueen, he was keeping me stylish. Now that’s
real, real, real, gun in my purse, bitch, I
came dressed to kill. Who wanna go first? I had them
pushing daffodils. I’m high as hell, I only
took a’half a pill. I’m on some dumb shit

68
For one thing, it seems skewed to detach the pickup syllables preceding the downbeat of every bar from the syntactical unit to which they belong—that is, shifting each pickup down a line would make the transcription itself appear less confusing. In addition, the final three lines feature both syntactical closures and rhymes in the middle of the bar. At a tempo like the one in ‘Anaconda’, one alternative might be to transcribe the track in double time compared to the above example, but that would result in nonsensically short lines and fail to resolve the pickup issue as well.

Perhaps, in fact, lyric transcriptions are not the arena in which to represent the musical metrical structure. It appears that rappers themselves do not transcribe their lyrics in alignment with the musical metre. While I have not conducted a rigorous study of a large amount of data, I spot-checked random verse transcriptions by three different Norwegian rappers and found little overlap with the musical structure. For example, I opened ‘Gatas Parlament: røverhistorier og raptekster’ (Don, Borgersrud, & Borgersrud, 2009) at random and checked the first verse of the song on page 131 (Gatas Parlament – Ti Mot En), where the first four lines are transcribed as follows (2002, from 0:44):

Gatas Parlament, det er oss med de fengende refrengene,
så lenge det rocker, gir vi pokker i pengene
Vi vil ikke hore for å menge oss med de store,
hvis det er noen som skulle tro det

If one were to segment the lines following the musical metre, however, it would look very different (the first word/line is a pickup to the first bar):

Gatas
Parlament, det er oss med de fengende refrengene, så lenge det rocker, gir vi pokker i pengene Vi vil ikke hore for å menge oss med de store, hvis det er noen som skulle tro det

This transcription might represent the musical metre, but the segmentation is otherwise nonsensical, which in turn undermines the point Caplan and Pate make about the transformation of sounding rap to poetry on the page. We see that in the printed version of Gatas Parlament’s lyrics, the lines are organised by primary rhyme, creating an ‘end-rhyme’ structure in the lineation, even though those rhymes are clearly not placed consistently at or around the fourth beat of the musical bars—that is, the ‘ends’ of the musical metre.
The second lyric transcription I checked included the first verse of Norwegian rapper Krigsskip’s 2018 song ‘Fenomenal’ (a favourite of mine). Through personal communication, I know that Krigsskip himself posts his lyrics to genius.com. Interestingly, though Krigsskip rhymes only sparingly (and in a very idiosyncratic style), his chosen line segmentation still does not follow the musical metre. Nor does the third transcription I checked, of Side Brok’s ‘Mann med manera’ (2009, lyrics from https://sidebrok.no/tekster/ekte-menn/), where rapper Runar Gudnason renders his lyrics as two delivered lines per text line, separated by a slash. The first text line reads:

\[
\text{ej e en mann med manera / svarte og blonke sko matcha augene og sjela}
\]

Even allowing for pickups and syncopation, this segmentation produces three beats in the first line and five in the second. Gudnason evidently prefers segmenting by parameters other than musical metre—in this case, syntax and primary rhyme position.

While this informal sampling does not suggest musical metre as a relevant guiding parameter for lyric transcription, rappers do think about musical metre a lot. In fact, it is at the very core of their musical expression, as Gudnason eloquently expresses in an article in the Norwegian newspaper Dagbladet (2016, my translation):

Almost every rapper writes lyrics. You need pen and paper or keyboard and screen to remember rhymes, structure ideas, and compose finished lyrics. There are different ways to represent the lyrics, but in the end they get a physical form; a form that in some way has to have a relationship to the number four.

Whether you use slashes or other lines to separate lyrical lines and bars, whether you write one or two lyrical lines per line, whether you use pen or keyboard, lower- or uppercase letters, your lyrics get your form. It’s a form that many of us can visualize, often like some sort of diagram, and it can help us to remember lyrics on stage. When I do a close study of other rappers, I sometimes place them into my own diagram, or let them hover above.

Gudnason exemplifies this diagram with an excerpt of Scarface’s verse on Geto Boys’s ‘Mind Playing Tricks on Me’, as shown in figure A4.3:
This flow diagram bears close resemblance to the one Edwards introduces in *How to Rap* (2009, p. 68) via The Pharcyde’s ‘Drop’ (1995), shown in figure A4.4:

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Let me freak the more junk than rate that compli-masquerade.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>funk, obso- Sanford sells. cate their mental</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lete is the state as I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I jet pro- pel at a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>punk that talks invade their</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>clipper . . .</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Both diagrams are set up as a table wherein each line of lyrics spans one bar of music, and there are four vertical columns representing each beat of the 4/4 metre. Both Gudnason and Edwards also use bold type to mark which syllables appear on the beats. In their writings, both also emphasise the importance of the placement of pauses, and this is evident in the way in which they visually annotate pauses which correlate with the beat positions. In many ways, this is reminiscent of the sixteenth-note grid systems we looked at in track A2.

However, this system is not analytical but compositional in nature: ‘The flow diagram used in [How to Rap] was created based on the systems many artists use to write down the flow of their raps’ (Edwards, 2009, p. 67), and the layout of one bar per line (meaning line on the paper, not poetic line) makes intuitive sense when one envisions a rapper actually jotting down lyrics. A famous quote by Rakim from the documentary ‘Something from Nothing: The Art of Rap’ (2012, 20:36) describes a variation upon this process:

I start off with 16 dots on a paper, bam, bam, bam, bam, bam. My thing was, if four bars was this long, . . . I see like a graph in between them four bars, and within that, I could place so many words and so many syllables.
The process described by Rakim and the various rappers interviewed by Edwards is similar to my own, where an ad hoc system of underlining and numbering organises early drafts of a rap verse in a notebook or any readily available piece of paper (see figure A4.5).

Figure A4.5: One of the author’s notebooks, ca. 2013. Note the numbering of bars and underlining marking beats (a parallel to the bold type in Gudnason and Edwards’s flow diagrams).

The principal affordance of this system is that it allows for flexibility when one is creating, changing and adapting a flow. No specific rhythmic values are indicated, and the underlining and numbering are partly mnemonic devices (for remembering approximate rhythms) and partly a way of making sure everything fits within the boundaries of the chosen structure (in the example in figure A4.5, the common sixteen-bar structure is suspended and extended by an additional bar, which turned into two additional bars in the final recording of the track). The system is, by design, flexible and sparse; it is also more prescriptive than descriptive, and therefore more useful to a performer than to an analyst. A straightforward flow diagram is how a rapper would answer the question ‘how would I perform this?’ (see the final sentence in the Gudnason quote above). That question, of course, is significantly different from ‘what is the structure of this flow?’ or ‘what makes this dope?’. 
While analysis has many things in common with the compositional process (as discussed in track A1), and the relevance of that process, from conceptualisation via production to sounding realisation, cannot be overstated, the flow diagram has some clear limitations as an analytical tool. Like the lyric transcription which slavishly follows musical metre, the flow diagram’s graphic layout obscures the linguistic and poetic structure of the lyrics. Some variations upon it attempt to rectify this by highlighting rhymes, for example, but this does not change the fact that the flow diagram overemphasises the musical metre in relation to any other structuring parameters.

**Lineation: What is a line?**

What distinguishes all poetry from prose is that poetry is made up of lines (verses). Syllables, words, phrases, clauses and sentences are found in both prose and poetry, but only poetry has lines. It is the organization of the texts into lines that defines poetry in all languages and literary traditions. (Fabb, Halle, & Piera, 2008, p. 1)

It seems like a simple, straightforward organising concept—the division of text into lines. We are used to reading poetry, seeing this text on paper or screen, graphically divided into lines. Lines, and the act of dividing a text into them—called *lineation*—is an explicit feature of any poetic text. Like most of the structuring concepts explored in this thesis, however, lineation is more complex than it might at first appear, and rap flow is a very fitting poetic expression through which to explore this complexity.

In *Language and Literary Structure*, Nigel Fabb argues that lineation is an implied form, not an inherent fact of the text (Fabb, 2002, p. 136). He allows for many kinds of lineation evidence which suggest alternatives regarding how a text can be divided and insists that the presence of these competing potential lineations is integral to the aesthetic expression of poetry. A central point here is that the graphic line break is only one type of lineation evidence, and that ‘inferences about lineation need more than one type of evidence’ (Fabb, 2002, p. 137). While this might seem counterintuitive (after all, the graphic line is often just referred to as ‘a line’), it does address the fact that poetry can exist in forms other than conventional printed verse (and rap is one of them), and that even in those conventional

---

63 As Agawu points out, analysis is like both performance and composition. He offers a snapshot of an analytical process: ‘We push forward in a compositional mode, playing with elements, rearranging them to see what might have been, and entering into rigorous speculation about music as intentional discourse’ (Agawu, 2004, p. 274). Interestingly, the scholar whose work Agawu polemicses against—Kerman—seems to be in agreement, noting that ‘the more fundamental alignment of music theory is with musical composition: for if theorists have an intellectual interest in the structure of music, composers have this same interest from their own strictly practical viewpoint’ (Kerman, 1985, p. 14).
contexts, the poet’s choice of graphic line breaks is never arbitrary but rather in some sort of relation with other lineation evidence. *The line* is not a graphic organisational unit but an experienced formal unit—even some printed poetry, that is, might allow for experienced lines which differ from the printed ones, depending upon the impact of certain other organising principles of text such as rhyme or linguistic (that is, syntactical or phonological) boundaries.

Since any text is subject to a variety of coexisting and/or competing forms of lineation evidence, it will typically enable or at least allow for several potential lineations.

In most texts, one lineation is dominant but I will suggest that other alternative lineations remain weakly present. ( . . . ) I propose that the weak implicatures of alternative lineations are experienced as aesthetic. In some texts, no single lineation is dominant, and the text is ambiguous in lineation, which is a kind of high-level complexity, which in turn may be experienced as aesthetic. Thus verse is inherently contradictory and complex and thereby inherently aesthetic. 64 (Fabb, 2002, p. 136)

This means that when one is analysing poetic metrical spans in rap flows, one must determine what the dominant lineation is (if there is one) while also accounting for whatever weak alternative lineations are present, as well as their impact upon the experience of the structure of the rap flow. Pivoting back to Bradley, one can interpret his desire to formalise the line/bar coincidence in rap lyric transcription as a way of generating a simple rule equating musical metre with dominant lineation, as one could be tempted to do with the graphic line in printed poetry. However, the graphic line does not always coincide with the dominant lineation, though it tends to be a fairly sturdy type of lineation evidence. Likewise, musical metre does not always coincide with the dominant lineation in rap flows either, and I will argue throughout this thesis that in this case it is not even a particularly sturdy type of lineation evidence.

As previously discussed, several rap scholars have analysed lines or line-equivalent spans, but they are either looking at a range of lineation evidence or they use the term ‘line’ without accounting for its nuances. Both Condit-Schultz and Ohriner’s corpus studies gather data on specific lineation evidence, including rhyme position, linguistic (syntactical and/or phonological) boundaries and breathing pauses in the rappers’ delivery. The analysis of ‘lines’ is a different matter, however, and Ohriner’s application of the term in ‘Metric Ambiguity’ (2016a) is useful there but ill-suited to an analysis of lineation as an implied form. There is no

64 Note that Fabb uses ‘verse’ in the sense of text divided into lines.
reason for the number of musical bars and poetic lines to always be the same, even if it is a common occurrence.

Other scholars employ the term ‘line’ while discussing parameters of the organisation of rap flows. In practice, an analysis of interactions among various types of lineation evidence. Katz’s (2008) analysed corpus consists of ‘1097 lines of hip-hop’ derived from thirteen songs by various artists, indicating his confidence in an objective definition of line. In a later publication, he remarks that the line is a constituent which is typically defined through certain types of lineation evidence (without using this particular terminology), and that line endings typically align with other parameters like rhyme and musical metre. (Katz, 2015a, pp. 9-11). Katz considers line endings to be defined by linguistic boundaries, and he observes that the line/bar coincidence is not universal:

Although musical rhythmic units [i.e., bars of musical metre] tend to align with linguistic constituents [i.e., line endings], mismatch between the two types of constituent is fairly frequent as well. Although rhymes generally do occur at some more or less predictable rhythmic interval, they are not constrained to appear only in this position. (Katz, 2015b)

Adams (2020) lays out a theory which emphasises the importance of some types of lineation evidence without employing the term ‘line’ as such. He analyses deviations from ‘rhymed couplets with end-rhymes falling on beat four’ (that is, the line/bar coincidence) and calls them ‘phrases’. Like Katz, Adams stresses the power of linguistic constituents (specifically ‘lyrical syntax’) as lineation evidence. In addition, he presents harmonic movement in the musical background and the repetition of rhythmic motives as parameters which can define ‘phrases’.

Another type of lineation evidence that is central to most analyses of rap flows is rhyme. Pate goes so far as to suggest that lineation is implied, and that rhyme is strong evidence for it: ‘Most times, we can sense line breaks (…) by observing the rhyme’. He then continues, ‘I encourage the readers/listeners who want to investigate the construction of rap/poems to focus more on the rhyme patterns as opposed to the metrical and syllabic computations we use in scansion’ (Pate, 2010, p. 111). Komaniecki finds that line and musical metre can be readily

---

65 Note that the preprint of ‘Hip-Hop Rhymes Mirror Phonological Typology’ published by the author online (available at https://community.wvu.edu/~jokatz/LenitionNonAnon.pdf) differs slightly from the published article. The passage cited here is from the preprint version.
detached from one another, and that rhyme is a strong type of lineation evidence, through his analysis of different ‘metrical positions of end rhymes’ (Komaniecki, 2019, p. 80). The entire concept of ‘end rhymes’ indicates that they appear at the end of a poetic constituent (that is, a line); because end rhymes appear in other metrical positions as well, the line/bar coincidence must necessarily be a convention, not a rule.

To summarise: rap is a poetic form, which means it has lines. Or rather, rap has lines, which means it is a form of poetry. Lineation is an implied form, and it is implied by the triangulation of different types of lineation evidence. There are many types of lineation evidence, but the ones which have been identified as the most central by rap scholars are linguistic constituents (that is, syntactical and phonological units) and rhyme. Additionally, musical metrical spans—bars—and their tendency to coincide with lines (the line/bar coincidence) have received significant scholarly attention as an organising formal unit of rap flows. I will return to musical metre’s relationship with lineation later in this track, but first, a slight theoretical detour.

**Is rap ‘metrical’?**

When I call lines ‘metrical spans’, this is, in fact, a misnomer. Lines do not necessarily imply poetic metre, and poetic metre, as we have seen, is only one of the many types of lineation evidence. My use of ‘metrical span’ rather than the more technically correct ‘lineation span’ or even simply ‘poetic span’ is rooted in the analytical dimension of metre, both musical and poetic, as I described in the first part of this track (see also figure A4.1). The parallels between pulse-/beat-level units, subdivisions of said units and spans of grouped units indicate different levels of (potential) interaction between musical and poetic structure. However, one can and should question whether rap is ‘metrical’ in the poetic sense, and if it is (or is not), whether this matters.

As previously discussed, poetic metre and metricality can be thought of as a meeting of ‘actual prose rhythm’ (the explicit rhythmic performance) and ‘silent metrical rhythm’ (an idealised schematic pattern). For poetry to be metrical, there must be some sort of pattern to the rhythm of the words (‘rhythm’ meaning, in this context, the syllables and/or stresses), and this rhythmic pattern must fit with a type of either pre-existing or emergent reference structure. The archetypical example of such a reference structure would be iambic pentameter—groups of five iambics, or feet consisting of two syllables, wherein the first is unstressed and the second stressed.
Rap flows can be poetically metrical, in the sense that they can consist of repeated spans of more or less identical stress patterns. However, they are not always metrical in this sense, and there is no great contrast between flows which are metrical and those which are not, or between metrical passages and passages which do not display consistent metrical patterning ('metrical’ in this passage meaning poetically metrical). This is because rap flows are structured according to musical as well as prosodic rhythm—they are performed alongside (or ‘on’) a musical background, or they invoke some musical rhythmic and metrical framework. Rap flow beat subdivisions tend to recall the canonical isochronous units of eighth and sixteenth notes and triplets (and various swung versions of these categories), and rap flow patterning displays some sort of musical phrasing. In short, a fruitful way to describe rap’s poetic rhythm is to say that the prosodic rhythm is fitted onto a musical rhythmic framework, and that any poetic metricality (that is, the grouping of syllables into feet, and the organising of these feet into some sort of pattern) is necessarily suppressed. The musical metrical framework carries with it different affordances than a completely free poetic framework, and something which might be considered to display ‘the same’ poetic metre (or prosodic reference structure) can also display vastly different musical rhythmic structures—see, for example, the dactylic structures in figure A4.6.

![Figure A4.6: Different music-rhythmic realisations of repeated dactylic feet.](image)

While these examples reflect the same basic prosodic rhythmic structure—dactylics consisting of a stressed and two unstressed syllables—they are clearly very different from one another in a music-rhythmic sense. In the top and bottom examples, every syllable is of equal categorical duration, but the rhythmic feels of the examples are completely different—the top one is a regular and insistent reinforcement of the pulse, while the bottom one is a cross-rhythm which challenges the pulse’s role as the timekeeper. Prosodic stress is clearly important to rap flows,
but its role there is unlike its role in most other poetic forms. Rap flows can display poetic metricality, and it can function as lineation evidence, but it is not an important structuring concept.

**Convergent and divergent metrical structure**

*I can trace my technical-analytic interest in the relationship between bars and lines to specific moments in my own compositional practice. One in particular stands out. I was sitting on a plane, trying to get the flow right on a verse I was working on (it was probably quite entertaining for those who glanced in my direction, as I was silently mouthing the words and gesturing with my hands, as one does while rapping), and it was a real struggle to line up the uneven lines with the bars. I ended up doing lots of counting, drawing and writing—even notating the entire verse in music notation—to finally piece the flow’s structure together. The manipulation of the relationship between lines and bars has become a central part of my rhythmic expression as a rapper.*

Experiences like this one have led me to a theoretical framework for analysing the relationship between musical metrical spans (bars) and poetic lineation spans (lines) in rap flows which is based on whether or not the flow conforms to the line/bar coincidence. When the bars and lines coincide, I call it a *convergent metrical structure*, and when they do not, I call it a *divergent metrical structure*. While I have argued that a rap flow’s poetic metricality at the beat (that is, feet) and subdivision levels is not a particularly significant feature, the relationship between the span-level lines and bars is significant indeed. Thus, when I use the term *metrical structure*, I am referring specifically to the relationship between bars and lines while disregarding the specific interactions between stressed syllables and musical beats. The musical background and the archetypical invoked musical metre in a cappella rap are almost always 4/4, and any rhythmic tension and variation at the metrical span level in a rap flow are created by lineation (or competing lineation evidence) which suggests alternative metrical groupings. Just as there is a dominant lineation alongside other weaker lineations, rap flows with a divergent metrical structure will (most times) have a time signature of 4/4 alongside other potential metrical groupings (suggested by the lineation) which are weakly present. As will be clear from the analyses on side B of this thesis, techniques based on different ways of creating a divergent metrical structure are prominent parts of rappers’ rhythmic arsenals, and the different types and degrees of divergence which are employed can be signature aspects of rappers’ rhythmic styles of expression.
There are potentially many types and degrees of divergent metrical structure, depending upon (1) the total volume or amount of divergence, and (2) the presence or lack of consistency or repetition of competing rhythmic patterns. In theory, then, the degree of divergence can range from minimal to complete, where minimal is a single point of divergence with fully convergent bars/lines surrounding it, and complete is a total lack of convergence of bar and line boundaries. The latter is, as far as I know, an entirely theoretical notion, as even in flows which display extensive divergent metrical structures, there tends to be at least some convergence, particularly at the end and/or beginning of a flow. We must also acknowledge the cases where the dominant lineation is challenged but not fully overpowered by alternative lineations, which results in a convergent metrical structure with a tendency towards divergence. This type of ambiguity allows a rapper to conform to the predictability of the line/bar coincidence while still creating some sort of structural tension at the level of the metrical span. We will explore examples of this borderline situation in track B1.

My inspiration for the terminology of convergence and divergence is the way Reuven Tsur uses it in his theory of cognitive poetics. Tsur roots his thinking in gestalt theory and applies the terminology of convergent and divergent poetry to the poetic structure’s convergence and divergence with ‘strong shapes’ (or gestalts). For him, gestalt theory is the most comprehensive theory that accounts most consistently for the relationships between structures and perceptual qualities or emotional qualities. Its application in art criticism in general, and in literary theory in particular, is one of the most illuminating ways to relate structures with effects. (Tsur, 2008, p. 112)

He connects specific aesthetic qualities with convergent (‘certainty and control’) and divergent (‘uncertainty or flexibility’) poetry and describes the two different styles:

[The structure of the] ‘convergent’ style is marked by clear-cut shapes, both in content and structure; it is inclined towards definite directions and clear contrasts (prosodic or semantic). . . . From the structural point of view, ‘divergent’ style is marked by blurred shapes, both in content and structure; it exhibits general tendencies (rather than definite directions) and blurred contrasts (prosodic or semantic). . . . The two are not solid categories; the differences are of degree, shades are gradual, along a spectrum. (Tsur, 2008, pp. 84-85)

Some of these descriptions apply to my concept of convergent and divergent metrical structure as well, particularly in terms of the ‘blurred contrasts’ and the gradual spectrum
from one to the other extreme. Significant differences also arise. Since a divergent metrical structure is the result of two metrical spans which diverge from one another (it is a specific structural feature), whereas divergent poetry describes a more general quality of the entire poetic structure, there is nothing stopping a rap flow from having a divergent metrical structure while being what Tsur would consider convergent poetry (that is, displaying clear-cut shapes and contrasts), as in the example in figure A4.7.

Figure A4.7: A divergent metrical structure with consistent repeated rhythmic groupings in Dizzee Rascal’s ‘Dirtee Cash’ (2011, 0:48–0:57).

In the opening of the first verse of ‘Dirtee Cash’, British rapper Dizzee Rascal employs an extensive divergent metrical structure with a fully consistent internal structure which displays the ‘clear-cut shapes’ indicative of Tsur’s convergent poetry. The presented lineation evidence—primarily the syntactic groupings and the rhyme position—indicates that the lines display a consistent cross-rhythmic (or even ‘polymetric’) structure spanning groups of three beats.

Note, however, that the third group of three beats consists of two subgroups spanning one and two beats, respectively, as defined by the same lineation evidence (linguistic structure and rhyme). Here, as well, there are other types of lineation evidence which might imply that ‘Aimless. People act shameless’ should be considered a line—namely, the individual line length (which might be expressed as ‘a line has to be of a certain length’) and the correspondence/parallelism between lines (which might be expressed as ‘if all surrounding lines are of a certain length, a line is probably of that same length’). The interpretation that a consistent three-beat-spanning lineation is dominant is an example of the structuring power of ‘strong shapes’ or gestalts from Tsur’s convergent poetry, and the way in which they can be used to triangulate lineation evidence.

I am not the first scholar to analyse the aesthetic qualities of various degrees of divergent metrical structure (even if I am the first to use that specific terminology). Mattessich (2019) explores the concept through three very different flows by Kendrick Lamar using the terminology of ‘derivative’ and ‘generative’ flows, which correspond to convergent and divergent metrical structures, respectively. While our theoretical concepts align quite well,
there are nuances which differentiate our approaches. First, I am hesitant to adopt the
terminology itself, as the concept of ‘derivative flow’ is based on the metrical structure being
‘derived from’ the musical background/surroundings. While I believe that to be a perfectly
valid stance to take, I discern a relevant if subtle difference between the derivation of
convergent metrical structure (or ‘derivative flow’) from its paired musical context and its
derivation from a historical convention which emerges from how rap flows have previously
been organised. Second, I take issue with the claim that ‘generative flow (…) functions
independently of the instrumental track (generating its own structure and internal rhythmic
relationships)’ (Mattessich, 2019), which is at odds with my own approach to divergent
metrical structure. In many (perhaps even most) cases, passages displaying a divergent
metrical structure generate their rhythmic effects specifically through their relationship with
the musical metrical framework as expressed in the instrumental track, rather than through
their own internal organisation. One could argue that a divergent metrical structure is often
more dependent on its musical surroundings than a convergent metrical structure is—its
rhythmic effects thus being more ‘derivative’. I also focus more on the concept of lineation
than Mattessich does, and specifically on the ambiguity of weak alternative lineations and the
interaction between different types of lineation evidence.

Returning to the zoom levels for analysis discussed in tracks A2 and A3, we can see that the
analysis of convergent and divergent metrical structure happens at the macrorhythmic level,
since the relevant structural units are typically around a bar long, and there is very little need
to look at rhythmic units smaller than a beat. Tracks B1 and B2 will both explore techniques
for manipulating the relationship between lines and bars, while track A6 will further flesh out
the common structural framework within which the lines and bars operate.
Track A5: Rhyme and its rhythmic roles

Rhymes are at the root of everything rap. One can probably make the argument that rap as a musical expression found its form when the first MCs started rhyming, rather than talking or shouting, over the DJ’s cuts. Edwards puts it simply:

"Rhyme is often thought to be the most important factor in rap writing—MCs often refer to rap lyrics as ‘rhymes’. Along with rhythm, rhyme is what gives rap lyrics their musicality, because similar sounds being repeated are interesting to listen to."

(Edwards, 2009, p. 81)

Scholars such as Pate (2010) refer to rap verses as ‘rap poems’, and there is a widespread tradition of separating the lyrics from the music and conducting a purely literary analysis of rap verses. Other literary scholars have written entire books where rhyme is either the sole focus (Caplan, 2014), one of the main topics (Bradley, 2009) or a part of the title, even if the focus is more on hip-hop culture than rhymes as such (Williams, 2013). It is also an aspect of this thesis’s definition of ‘flow’: flow is the rhythm of the words and rhymes in a piece of rap music.

Rhyme, fundamentally, is speech-sound parallelism—that is, the repetition of some sort of speech-sound within a relatively short time span. However, not all speech-sound parallelism is rhyme, and different traditions, disciplines and art forms have different criteria for what constitutes ‘rhyming’. For one thing, alliteration—words beginning with the same consonant sound—is a different type of speech-sound parallelism than rhyme, and it has a different effect. This goes for other types of consonance—the repetition of consonant sounds—as well. Assonance—the repetition of vocal sounds—are indeed typically treated as full-fledged rhymes in a hip-hop context, whereas some traditions reserve the term ‘rhyme’ for what is otherwise often called ‘perfect rhyme’, wherein all the sounds except the initial consonant sound are repeated. Extensive analyses of the aesthetics of rhyme delve into the typological nuances of perfect rhymes and assonance, as well as various descriptions of the rhymes which are not perfect but not mere assonance either, using terms like ‘slant rhymes’, ‘partial rhymes’ and ‘half rhymes’. When rhymes stretch over many syllables (multi-syllable rhymes, often

66 I use ‘speech-sound’ rather than ‘phonological’ or ‘phonetic’ because the latter terms impart certain nuances, depending on one’s interpretation of them. ‘Phonological’ implies that the sounds are semiotic (i.e., ‘language sounds’), whereas ‘phonetic’ points to sounds made by mouths whether they are semiotic or not. This divide is really not as clear cut as these definitions might imply, however (Blankenship, 2021, pp. 11, fn. 13).

67 See Fabb (2022 [forthcoming]).
colloquially referred to as ‘multis’ when they consist of three or more syllables), and sometimes across several different words (‘mosaic rhymes’), the aesthetic effect is very different from that associated with simple monosyllabic rhymes. Likewise, rhymes which are the result of a rapper’s exaggerated or altered pronunciation of a word will stand out more than a more conventional (or, worst case, the dreaded predictable) rhyme.

A novel approach to the analysis of rhyme is that of Michael Blankenship, whose dissertation is devoted to creating ‘a system for organizing and conveying the various aspects of speech sounds as music’ (Blankenship, 2021, p. 4). With this system for the transcription and representation of speech sounds, reminiscent of music notation, Blankenship intends to elevate the status of rhyme (and speech sounds in general) as a musical parameter. Identifying ‘rhyme’ is not enough, according to Blankenship; the interaction of repeated identical and similar speech sounds that constitutes rhyming must be transcribed, represented and analysed to show how these sounds rhyme, not merely that they do.

In an analysis of rap, referring to sequences of syllables that bear a wide range of similarities and differences in phonemic content, stress pattern, cardinality, rhythm, and metrical location simply as ‘rhymes’ is like an analysis of tonal music in which harmonies are labelled arbitrarily according to order of appearance (Chord 1, Chord 2, etc.), and the only harmonic relationships identified were that of congruency between any chord progressions featuring similar root motion, regardless of tonal syntax. (Blankenship, 2021, p. 231)

I find Blankenship’s goal of exploring and highlighting the nuances of the speech sounds of rap as music to be admirable, interesting and practical, though some of his claims and justifications are problematic. The above quote, for instance, is less an argument against analysing rhyme as a distinct experiential category than an argument for an advanced system of labelling different rhyme instances in a particular manner regardless of whether or not the analysis at hand requires it. The comparison to harmonic progression is thin as well. While rhyme and its development can encompass both variation and nuance, rhyme does not operate like functional harmony, whose development features highly formalised (genre-dependent) patterns of interaction and ordering. There is most definitely room for broader scholarship on the intricacies of speech sounds and rhyme in a musical context, and the system Blankenship has developed seems well suited to such an endeavour. However, if we view rhyme as a category with both digital and analogue dimensions (as was discussed in track A3), the
experience of something rhyming (its digital quality) can in principle be analysed without taking into account the nuance of how it rhymes (its analogue qualities).\footnote{To Blankenship’s credit, he does demonstrate how his system can be used to show that something may well rhyme even if it is not immediately apparent that it does. See, for example, his comparison of his own and Komaniecki’s analyses of Cardi B’s 2018 track ‘Bodak Yellow’ (Blankenship, 2021, pp. 282-286).}

This thesis reserves itself to rhyme’s rhythmic roles and spends relatively little time on rhyme types—that is, how something rhymes, whether its phonological connections are stronger or weaker, and so on.\footnote{For a discussion of the different typologies of rhyme, see Sykäri and Fabb (2022 [forthcoming])—for example, a ‘theoretical typology’ is the reasoning behind the treatment of alliteration as a uniquely functioning speech-sound parallelism other than rhyme (Sykäri & Fabb, 2022 [forthcoming]).} Like most rap analysts, but unlike Blankenship, I will treat rhyme as a binary at the categorical level—either something rhymes, or it does not. And if something rhymes, the question is most often what the rhyme’s role in the rhythmic and formal structure of the rap is. Deciding what rhymes and what does not is another example of something that seems intuitively simple from a subjective experiential point of view but is rather difficult to pin down objectively and empirically (as was the case with ‘rhythm’ in track A1).\footnote{Blankenship challenges this claim, insisting that his method is ‘a method for transcribing lyrics capable of capturing their phonemic and musical content without being influenced by any particular analytical purpose, so those visual representations can serve as a mutually agreed upon set of basic facts about a song, from which a more epistemologically transparent analysis can proceed. In other words, we need transcriptions that are more like a traditional musical score’ (Blankenship, 2021, pp. 233-234). I do not believe his approach (or any other positivistic approach, for that matter) is capable of producing ‘basic facts about a song’, but I applaud his ambitions in this regard.} Several rap analysts have stressed this point:

> When identifying rhymes (. . .) there were many circumstances where I was unable to state with absolute certainty whether or not a pair of syllables rhymed. (. . .) Readers should bear in mind that my analysis of rhyme in hip-hop will reflect the criteria outlined above, but also be influenced by an unavoidable portion of subjectivity. (Komaniecki, 2019, p. 72)

In actual artistic contexts, identifying when a rhyme occurs can be somewhat subjective. This is especially the case with singleton rhyme motives, such as a single vowel or consonant. Since there are a limited number of vowels, utterances will frequently repeat vowels without evoking the qualia of rhyme. Only if a vowel is used often enough, or regularly enough, does it start to sound like a rhyme. (Condit-Schultz, 2016c, p. 132)
One must identify which time points listeners experience as rhymes. As I will show, this cannot be formalized in the same manner as accent (...) listeners seem rather permissive in judging [rhyme]. (Ohriner, 2019b, pp. 64-65)

The sheer subjectivity of the experience of rhyme has consequences for any analysis of rhyme in rap, since the scholar is subject to it as well. The analytical work of repetition, transcription, visualisation and so on might result in the ‘appearance’ of rhymes in the second, third or forty-seventh listening when they were not there, experientially, in the first. Interestingly, ERP research has shown that the experience of rhyme is impacted by (if not completely dependent upon) selective attention in any case, so there are probably no true ‘objective criteria’ for experienced rhyme. The perhaps most successful foray into the empirical evaluation of rhyme takes the subjective dimension into account—Hirjee and Brown (2010) uses an automated system wherein potentially rhyming syllables are given a score related to how likely they are to be experienced as rhymed. In the analyses in the present thesis, rhymes are annotated at the analyst’s discretion. When there is significant doubt as to whether a rhyme will be experienced as such by most listeners, the analysis accommodates this fact.

An important point to note is that rhyme in rap flows is not merely a speech-sound parallelism; most of the time, a rhythmic parallelism accompanies the phonological/phonetic one, so the rhymes rhyme rhythmically. In any polysyllabic rhyme, all rhyming instances will tend to display the exact same rhythmic figure. As Komaniecki notes, a mismatch is more likely in the speech-sound parallelism than in the rhythmic one: ‘While it is common in rap for repeated rhythmic motives to not strictly rhyme every syllable, the opposite—that is, polysyllabic rhyming pairs that don’t correspond rhythmically—is exceedingly rare’ (2019, p. 46). This is why it seems so effortless for rappers to make rhyme types which are traditionally considered ‘weak’, such as assonance, come across quite clearly as rhyme. This dimension of rhyme in rap will come up throughout the analyses in this thesis and should be kept in mind even when it is not always apparent in the representation at hand. A very clear example of this which showcases the stacking of types of parallelism in addition to speech-sound and rhythmic is the analytical amuse-bouche towards the end of track A6.

---

71 ERP, standing for ‘event-related potential’, is the measured electrophysiological response in the brain to some sort of event. See Praamstra and Stegeman (1993) and Yoncheva, Maurer, Zevin, and McCandliss (2013) for ERP research on rhyme.
Rhymes as boundary markers: *Primary rhymes*

Rhyme and other parallelisms’ rhythmic effects in a musical stream have a peculiar dimension to them, in that, when they appear, they modulate both the immediate present and the recent past. Initially, of course, a listener cannot yet know that the syllable or syllables in question will be a rhyme—the parallelism only acquires its structuring power upon the arrival of the second instance of the parallel event. This event, in turn, shares its rhythmic effect with both the immediate perceptual present and the preceding rhythmic structures (by informing and altering them). Of course, there are conventions regarding the positions within both musical and poetic metrical structures (that is, bars and lines) in which rhymes will commonly appear, so the listener might well suspect a rhyme, but there is no guarantee. The main rhythmically structuring role of rhyme in a rap flow, then, is to influence the joint metrical structure by either reinforcing or weakening the dominant lineation. The way it does so is by serving as a boundary marker.

As discussed in the previous track, rhyme is a sturdy type of lineation evidence, meaning that it often correlates with a line ending. This is why it has seemed fitting to import the concept of ‘end rhyme’ from the traditional literary analysis of poetry into the analysis of rap. However, I will argue that rap flows are (1) more likely than other forms of poetry to have rhymes in positions other than the end of lines, and (2) more likely to display significant ambiguity in their lineation. Due to the prevalence of convergent metrical structure—that is, the correspondence between musical metre and lines—the term ‘end rhyme’ can be misinterpreted as referring to the end of the musical metre, as evident in Bradley (2009, p. 42): ‘The most common rap rhymes are end rhymes, those rhymes that fall on the last beat of the musical [bar], signaling the end of the poetic line’. Another common feature of rap flows is that there are often more than two instances of each rhyme class which are strung together as part of a larger *rhyme complex*, a term coined by Krims to describe ‘a section of a song in which any one rhyme predominates’ (2000, p. 43). Rhyme complexes are another type of structural unit in rap flows whose boundaries will often, but not always, coincide with bar and/or line endings. Since a rhyme ‘predominating’ thusly need not be placed at the line endings, it becomes clear that rhymes other than ‘end rhymes’ can have significant structuring roles. In my analyses, then, I will refer to the most structurally significant rhymes—those that dictate the rhyme complexes—as *primary rhymes*.

A rhyme complex can be but a single line where two or more instances of the same rhyme class create an internal parallelism, but the line is not otherwise connected to another. Or a
rhyme complex can swallow an entire verse (or even several verses or an entire track) in cases of *extended monorhyme*, a term coined by Komaniecki (2019, p. 99) to label long sections or entire verses dominated by a single rhyme class. In this thesis, several analysed tracks showcase extended monorhyme, including Ms. Lauryn Hill’s ‘Doo Wop (That Thing)’ (1998) in track B1 and Chance the Rapper’s ‘Do You Remember’ (2019) in track B3. While extended monorhyme is a conceptual idea which has appeared in hip-hop since its early days, trends toward ever longer rhyme complexes and/or busier rhyme density have developed over time. Condit-Schultz notes that ‘the usage of longer rhyme chains became much more popular around 1998’ (2016c, p. 137), where ‘rhyme chains’ refer to all the instances of a rhyme class which are connected (that is, all [primary] rhyme instances in a rhyme complex). As will be evident throughout this thesis, the positions of primary rhyme instances within rhyme complexes vary considerably—they occur at the ends and beginnings and middles of lines; they converge with and diverge from bar and line boundaries; and they are occasionally ambiguous in their presentation, shifting gradually from one rhyme class to another.

**Additional prominence from additional rhymes: Secondary rhymes**

Even if rhyme is treated as a binary—either something rhymes or it does not—not all rhymes are equal. Or, more precisely, some rhymes are more equal than others, like the primary rhymes which dictate rhyme complexes and function as lineation evidence. But rappers do not only use rhymes which contribute to formal or metrical structure. They will often sprinkle additional rhymes throughout their flows, adding extra parallelisms and rhyme connections within and between the more dominant structuring units. Occasionally, flows can be so saturated with rhymes that it seems as though, in short, extra rhymes mean extra dope. And while an increased rhyme density can be achieved through more primary rhyme instances, one can also insert different rhyme classes as well, assuming they are not strong enough to wrestle control over the rhyme complex from the primary rhyme. I call all rhymes which are not primary secondary rhymes.

The analysis of rhymes of secondary structural importance has tended to be tied to the concept of the end rhyme in the analysis of both written poetry and rap. The dichotomy between ‘end rhyme’ and ‘internal rhyme’, or a rhyme within a line rather than at the end (and thus not serving to link two or more lines), also appears frequently. In rap analysis, it is central to Alim’s (2003) analyses of Pharoahe Monch and to Bradley’s summary of rap’s rhyme development over time (2009, pp. 42, 62-63). Of course, this again excludes certain types of non-primary rhymes, as not all secondary rhymes are internal to a poetic line. The reason to
avoid ‘end rhyme’ in analysis can thus be extended to ‘internal rhyme’. Edwards (2009, pp. 103-104) uses the term ‘extra rhymes’ for all rhymes which are not end rhymes, which begins to broaden the terminological scope, but there is still a problematic dichotomy at its base, as noted by Ohriner: ‘In printed verse, any rhymes not at a line ending are internal. But rap lyrics are not printed; line endings are often ambiguous, making internal rhyme hard to define’ (Ohriner, 2019b, p. 111). Using ‘secondary rhymes’ for all non-primary rhymes allows for the analysis of less structurally significant rhymes which straddle line boundaries, as will be exemplified in track B1.

Alongside their ability to create rhyme connections (parallelisms) across structural units, emphasising or blurring boundaries along the way, the main rhythmic role of secondary rhymes is that they add prominence to the rhythmic events of which they are a part. Rhythmic events which rhyme will stand out slightly from the events surrounding them. This is the basis of the idea of poetic accent and rhyme’s contribution to the topography of rhythmic events, as will be discussed in the next track. Note that this contribution to prominence is something that is common to all rhyme (and even other types of speech-sound parallelism), so both primary rhyme and secondary rhyme have this effect. Another crucial point is that a rhyme’s function might not be entirely clear, and it is not always possible to determine whether a rhyme is primary or secondary. There will inevitably be cases where two rhyme complexes overlap slightly, and the last rhyme instances of the first complex and the first instances of the second might be weakened, for example. There are also cases of partial rhyming, rhyme shifts or pivot rhymes (see track B1 for an example), wherein the rhyme classes themselves might be ambiguous. Which rhyme complex, if any, is the rhyme part of? Is it a secondary or a primary rhyme? Or does it have a primary rhyme connection to some rhyme instances and a secondary rhyme connection to others? Rhyme is one of the most important structuring concepts in rap flows, and its effects can be both profound and subtle. Many of the techniques described throughout side B of this thesis are based upon and named after the ways in which they employ rhyme and rhyme position. Since there are many different terms and concepts which relate to rhyme, the final part of this track will be dedicated to some definitions, to which the reader can return as needed.

**A short rhyme dictionary**

The following terms are either ones I use in the thesis (some more than others) or ones which are ubiquitous in common discourse (whether scholarly, general or hip-hop centred) and warrant an explanation. Some of the terms are adopted and/or adapted from other scholars or
practitioners, some are my own inventions or slightly different takes on existing concepts, and some are actually less useful for rap analysis but still encountered there. The terms fit into broad categories—many are typological, while others concern form and structure, and some are technical—but I have chosen to sort them alphabetically here to accommodate the section’s usefulness as a mini-dictionary.

Assonance: A vowel-sound rhyme, meaning that any consonant sound after the rhyming vowel sound is different across the rhyme instances. This is often expressed as ‘CVC’, indicating that the non-bolded C’s (consonants) are dissimilar, while the bolded V (vowel) rhymes. In multisyllable rhymes, this can extend to consonants between vowels as well (for example, VCVC). The term is often used quite broadly to indicate some sort of difference in the consonant sounds between rhyme instances or within the rhyme class, without necessarily implying that all the consonant sounds are dissimilar.

Alliteration: See also consonance. A specific type of consonance (consonant-sound parallelism) wherein the same consonance sound is repeated across the beginnings of words (CVC). The two terms are often used interchangeably outside of technical literary analysis.

Bridge rhyme: A rhyme which connects (‘bridges’) two otherwise separate formal sections (such as line pairs or larger structural blocks). See Alim (2003, p. 75) and track B2.

Chain rhyme: A section of high-density rhyming using a single rhyme class which is typically line-internal or otherwise localised to a shorter formal section. See Alim (2003); Komaniecki uses the term ‘rhyme cluster’ instead (2019, p. 95). Not to be confused with rhyme chain.

Consonance: Consonant-sound parallelism. Unlike alliteration, consonance refers to any sort of marked (as in ‘standing out’) consonant repetition across or within words. A favourite example of mine is from Side Brok’s ‘Forstå Kar Me Kjome Frå, Del 1’ (2006, from 1:45):

‘Moltø bene moltebeir, ei mormor i marmor med mjuke hender’.72

End rhyme: Line-final rhyme, meaning that it is placed at the end of a poetic line irrespective of the musical metrical structure. See discussion in this track regarding why I prefer analysing primary rhymes rather than end rhymes.73 When I need to specify that a rhyme is at the end of a poetic line, I will use line-final rhyme.

---

72 Note that in the first clause of this example, the first syllable includes assonance as well (CVCV).
73 There is also an extended discussion of this in Oddekalv (2022 [forthcoming]).
Extended monorhyme: From Komaniecki (2019, p. 99). One rhyme class is the primary rhyme for an extended formal unit of flow—often an entire verse (that is, an entire section or verse is one large rhyme complex).

Half-line rhyme: This is a specific type of internal rhyme, where one rhyme instance is placed at the halfway point of the line, subdividing the line into two equal parts. The half-line rhyme may rhyme with the line-final rhyme or the preceding line.\(^74\)

Hidden rhyme: A rhyme which is either ‘weak’ or unmarked and thus probably not experienced as a rhyme by most listeners (outside of, for example, repeated listenings and/or close attention to rhyme itself). Reasons for a presentation as ‘hiding’ include a lack of rhythmic rhyme or proximity to a stronger rhyme class. Examples of this phenomenon appear in track B2.

Identical rhyme: The repetition of the exact same word/homophone (CVC, for instance). It is by itself a very weak type of rhyme, but it is often employed as part of a larger mosaic rhyme (like CVCV CVC). The latter type will rarely be referred to as ‘identical rhyme’, however.

Internal rhyme: Line-internal rhyme. Typically used to describe rhyme instances within a line which rhyme with the line-final (end) rhyme. As discussed in this track, I prefer to avoid using the term and instead focus on primary and secondary rhymes. A line-internal rhyme can be either primary or secondary, depending on its structuring role.

Mosaic rhyme: Multisyllable rhyme which crosses word boundaries, such as ‘mosaic rhyme—go say it, Brian’.

Multisyllable rhyme: Also ‘multi-rhyme’, ‘multi’ or ‘compound rhyme’ (Edwards, 2009, p. 87). A rhyme class which spans multiple syllables, like ‘madness’/’bad guess’ (a two-syllable rhyme) or ‘criminal’/’similar’ (a three-syllable rhyme). Note that, due to the proliferation of rap rhymes spanning at least two syllables, enthusiasts’ and/or practitioners’ discourse will at times only refer to rhymes of three or more syllables as ‘multis’.

Non-rhyme: The lack of rhyme where rhyme is strongly anticipated, meaning that there is an experience like rhyme in that the unrhyming word/syllables are marked anyway (and often even more marked than with an actual rhyme, as there has been a broken strong expectation). While there is a lack of speech-sound parallelism, a ‘non-rhyme’ will often feature rhythmic

\(^74\) A half-line rhyme which rhymes with the preceding line is not technically an ‘internal rhyme’, which is one reason why I prefer to avoid the term.
rhyme (that is, it will have the same rhythmic figure as the word it is expected to rhyme with), further reinforcing the feeling of ‘there should be a rhyme here’.

One-rhyme: A rhyme instance on beat 1 of the musical metre which is connected in some way to the previous bar. There are many different types and flavours of one-rhymes. Many are presented and analysed in track B1.

Partial rhyme: Similar to, but not exactly the same as, ‘slant rhyme’, ‘half rhyme’ or ‘imperfect rhyme’, which are terms for all rhymes which are not ‘perfect’ (meaning that both vowel and end consonant match: CVC). Instead, ‘partial rhyme’ refers to rhyme instances which rhyme only partially with other rhyme instances—in a chain of three rhymes, for example, one of the rhyme instances rhymes only partially with the two others. All three of them rhyme, but, strictly speaking, one of the rhyme instances has a different rhyme class than the others—in other words, the rhyme class which connects two of the rhyme instances is different from the one that connects the last rhyme instance to the first two. As in: ‘Father tried (to) marshall my partial rhyme’, where the second syllable of the first three-syllable rhyme (‘-ther’) does not rhyme with the second syllables of the other two (‘-shall’ and ‘-tial’). Partial rhyme connections can also be emphasised or even created by consonance and/or rhythmic rhyme (the latter is typically the case with split rhyme).

Pivot rhyme: A rhyme instance which rhymes partially with a preceding rhyme class and partially with a following rhyme class without rhyming fully with either one. Thus, it is also a type of bridge rhyme, as it connects two different rhyme complexes. See track B1 for examples.

Primary rhyme: A structurally significant rhyme class which governs a rhyme complex. See discussion in this track.

Rhyme chain: A group of rhyme instances connected by the same (or a similar, or a gradually changing) rhyme class. A rhyme chain will most often consist of a row of primary rhymes, although a rhyme chain of secondary rhymes is also possible. I use rhyme complex rather than rhyme chain in my analyses. The difference between the two is that rhyme chain refers to only the connected rhyme instances, whereas rhyme complex refers to the entire structural unit within which the rhyme chain operates.

Rhyme class: A rhyme class categorises the shared speech-sounds between two or more rhyme instances (or ‘what rhymes’ in a rhyme). In a two-syllable assonance rhyme, for
instance, the rhyme class consists of the two vowel sounds shared between the rhyme instances. Some scholars (Condit-Schultz, Blankenship) use the term *rhyme motive* instead.

**Rhyme connection**: A relationship between two or more rhyme instances. It generally derives from the fact that the rhyme instances share the same or similar rhyme classes (but not always—a non-rhyme can have a rhyme connection with whichever syllables it is expected to be rhymed with, for example).

**Rhyme complex**: ‘A section of a song in which any one rhyme predominates’ (Krims, 2000, p. 43). See also *rhyme chain* and discussion in this track.

**Rhyme density**: The amount of rhyme in the flow or a section of flow. Condit-Schultz measures rhyme density by ‘proportion of rhymed stressed syllables’ (2016c, p. 136). Komaniecki uses the related term *rhyme frequency* in addition to density, where frequency also takes into account the rhyme instances’ positions within the musical metrical and formal structure, which helps in the analysis of *rhyme regularity* (Komaniecki, 2019, pp. 72-83).

**Rhyme entropy**: A term employed by Condit-Schultz to measure the variation in distribution of rhyme position within the metre in a verse, track or larger corpus (see discussion in track A2). It is similar to Komaniecki’s *rhyme regularity*. I do not explicitly analyse this parameter in this thesis.

**Rhyme instance**: A single instance of rhyme (one or more syllables). A rhyme instance can be referred to without necessarily taking its connected rhyme instance(s) into account (it is, of course, implied that there is some sort of rhyme connection present). Often, ‘rhyme’ is used as a shorthand for rhyme instance if there is little potential for confusion/ambiguity.

**Rhyme position / Rhyme placement**: The position of a rhyme instance within the musical metrical structure. *Rhyme placement* is the ‘active’ term, acknowledging that a rapper places a rhyme instance at a specific position.

**Rhyme scheme**: A term for the structure of the rhyme positioning within a larger structural unit. The term is commonly applied in fan/practitioner discourse, as can be seen in part II, chapter 6, of Edwards (2009, pp. 95-110). Komaniecki uses *rhyme regularity* as a parameter when analysing rhyme schemes, whereas Blankenship (2021, p. Chapter 5) analyses regularity/irregularity using the terms ‘schematic’ (for regular) and ‘sporadic’ (for irregular) rhyme styles.
Rhyme shift: Coined by Komaniecki, rhyme shift is described as ‘a process where rappers perform a series of rhymes in which rhymed syllables are gradually altered over the course of an excerpt, leading listeners to hear an entire unit as being unified by rhyme even if the first and last rhyme in that unit don’t strictly rhyme’ (2019, pp. 70-71).

Rhythmic rhyme: The term used for the almost universal phenomenon of rhymes in rap corresponding rhythmically as well poetically (that is, phonetically/phonologically). The term comes from Komaniecki and is discussed both in this track and throughout the various analyses on side B (as well as in track A6).

Secondary rhyme: Discussed at length in this track, secondary rhymes are all non-primary rhymes, meaning that they do not govern a rhyme complex.

Split rhyme: A type of multisyllable rhyme where the first and last part rhyme but there are non-rhyming syllables between them (that is, it can be considered to be two separate rhyme classes rather than one compound one). This could be expressed as, for example, CVC-CVC-CVC-CVC. In split rhyme, the non-rhyming syllables will most often rhyme rhythmically, so the entire unit is experienced as rhyming. See example in track B1 (figure B1.5, p. 141).

Twisted rhyme: From Norwegian ‘vridd rim’ (Oddekalv, 2017, pp. 34-35). Also called ‘bent rhyme’ (Blankenship, 2021, p. 349; Komaniecki, 2019, p. 155). This term addresses the way in which rappers alter the common pronunciation of a word to create a rhyme where there normally would not be one.

The practice of twisting, bending, altering or mis-conjugating words for both rhyming and the sheer joy of playing with language is a favourite technique of the rappers in my group. The track ‘Legendisk’ from our 2020 album Gangsta Norvegicum is based upon this concept and how we in our internal lingo use the suffix ‘-isk’ in all kinds of words and contexts it does not belong to. The name of the album is also a result of a passage of mine from an older track called ‘Munch’ (Sinsenfist, 2017, from 1:30): ‘Det e ditt privilegium å sjå Gangsta Norvegicum—å bade i min aura e et ganske sweet opplegg i grunn’, where the stress pattern of the word ‘opplegg’ is altered from strong-weak to weak-strong to rhyme ‘-legg i grunn’ with ‘-legium’ and ‘-vegicum’ (and the preceding ‘delirium’ and ‘imperium’).

This is not even close to an exhaustive list of the terms and concepts which rappers, scholars, heads and poets use to capture and describe the vast phenomenon of rhyming. It is simply
intended to inform the reader about the ways in which this thesis engages with rhyme as a contributor to the rhythmic and formal structure of rap flows. The next track presents and explores the different levels of musical structure within which rap flows exist and which they are part of creating—levels wherein rhyme is a crucial piece of the intricate rhythmic puzzle.
Track A6: Levels of musical structure

The necessary theoretical background remaining ahead of the analyses of side B is related to the different structural levels of the rhythms of rap flows. There are some clear parallels to the zoom levels, and we will also revisit previously introduced concepts like metre, prominence/stress/accents, expressive timing and the digital versus the analogue, and the latter—that is, microrhythm’s contribution to a blurring of the boundaries between flow and delivery. Accompanying these theoretical discussions are discussions about the methodological implications of these concepts. This will culminate in a short analytical amuse-bouche and a presentation of my ‘methodological ideology’ regarding representation on the different zoom levels.

Groove and song modes of listening

One of rap’s spiritual and musical forebears is funk. Of the musical fragments used as ‘breaks’ by early DJs, sampled on both early MPCs and modern digital audio workstations (‘DAWs’) and reproduced faithfully by studio musicians in the ‘G Funk era’, passages from funk tracks have been recontextualised in hip-hop from its very inception. The most in-depth academic publication on funk and its musical poetics is Danielsen’s Presence and Pleasure (2006), and amongst its many insights which are applicable to a rap (and particularly hip-hop) context is the distinction between what Danielsen calls the ‘song mode’ and the ‘groove mode’ of listening.

The song mode of listening is the default position of most musical analysis and probably also of many pop/rock listeners in most listening situations, at least if we limit the scope of discussion to contemporary Western music. In a song mode of listening, we take the musical gestures of the immediate musical surface and attempt to organise them into larger structural pieces—verses, choruses, bridges, intros, outros and so on. When one is in the song mode of listening, the musical present becomes a point on the path towards a (possible) closure of some formal section in the music which will reveal whether our postulations about the formal structure were accurate or not. When this happens, we can pat ourselves on the back and/or

75 Described as ‘short passages of solo drumming that typically occur in funk songs’ (Oliver, 2015, p. 1), breaks, and the very concept of looping them, are clearly inextricably tied to funk music.
76 ‘MPCs’ are a series of famous drum machines made by AKAI.
77 ‘G-Funk’, short for Gangsta Funk, is a subgenre of hip-hop which originated in Los Angeles around the beginning of the 1990s. Pioneering producer Dr. Dre often employed live musicians to record new versions of funk grooves to avoid having to ‘clear samples’. The foremost avatar of the ‘G Funk Era’ is probably Warren G, whose debut album ‘Regulate . . . G Funk Era’ from 1994 was a huge commercial (and musical) success.
adjust our expectations as we raise the mental scaffolding for the next formal unit whilst eagerly awaiting its arrival.

The groove mode of listening, on the other hand, is about *moving forward while staying put*, or, as Danielsen describes it, the state of ‘being in funk’:

> [When the music invites a groove mode of listening,] our inclination as listeners to organize the musical material into an overall form gradually fades away. Instead of waiting for events to come, we are submerged in what is before us. Our focus turns inward, as if our sensibility for details, for timing inflections and tiny timbral nuances, is inversely proportional to musical variation on a larger scale. When funk is experienced in this way, music ceases to be an object that exists apart from us. The relation of subject and object is almost suspended. We operate within a continuous field where the limit between music and listener is not yet established or has vanished. Dancing, playing, and listening in such a state of being are not characterized by consideration or reflection, but rather by a presence in the here and now of the event. (Danielsen, 2006, p. 144)

Musically, what happens is that there is a repeated *basic unit* (Danielsen, 2006, p. 43), a musical tapestry (or composite weave, even) lasting one or two bars, which makes up the groove. This pattern is repeated without interruption and does not lend itself to the construction of larger formal units as such.

Danielsen stresses the importance of the physical aspect of ‘being in the funk’ and points out the dance floor as the arena where many (herself included) experienced the transcendence of a groove experience. In the studio and rehearsal space, the physical dimension is also central to the discourse within my own group. A programmed loop or played basic unit (which we typically refer to as ‘a groove’, fittingly enough) is considered promisingly groovy if the listener ‘får haud’—meaning that he starts bobbing his head in time with the music (‘å få haud’ literally means ‘to get head’ in my sunnmørs dialect of Norwegian—the figurative meaning of that phrase in English is not shared between the languages).

---

78 This is similar to the concept of ‘being-in-meter’ (1996)—an adaptation of the Heideggerian concept of ‘being-in-the-world’—which Steven Friedson uses to describe experiential dimensions of Tumbuka ‘vimbuza’ drumming of northern Malawi.

79 The ‘trance-like’ experience that different types of dance music can invoke in dancers and listeners is a recurring topic amongst dance-music researchers; see, for example, Becker-Blease (2004).
The meetings of the groove and song modes of listening and the beat and flow in rap form an area of floating boundaries, ambiguities and dualities. The repetition of short basic units which defy structuring into larger formal units can be in turn supported or contradicted by the interplay between musical background and lineation structure. As Danielsen is keen to stress, the two modes of listening can and will coincide, where some aspects of the music emphasise the repetition of the groove while others direct us towards organising the music into larger sections. An interesting example of these floating boundaries within a rap context is the veritable flagship of G-Funk—Warren G and Nate Dogg’s Grammy award–winning hit song ‘Regulate’ from 1994. The sample upon which the track is based is a four-bar loop from Michael McDonald’s ‘I Keep Forgettin’ (Every Time You’re Near)’ (1982). A periodicity of four bars is longer than the typical basic unit of a funk groove, but it can easily be thought of (and experienced) as two variations of a very similar basic unit because bars 1 and 3 are nearly identical. The four bars of the loop/basic unit converge perfectly with the way in which Warren G and Nate Dogg structure their sections of rapping. As can be seen in figure A6.1, the two rappers alternate whenever the loop loops, filling their four bars with a fully symmetrical convergent metrical structure with two line pairs (true ‘couplets’, in this case) connected by rhyme. This results in a likewise symmetrical hierarchy of periodicities which can be extrapolated all the way from a single bar to the entire thirty-two-bar period which runs from the intro to the instrumental interlude.

Figure A6.1: Symmetrical periodicities in ‘Regulate’. Fully convergent metrical structure, with every two lines (and no more!) connected by rhyme (the shortest periodicity here marked with ‘R’). The rappers alternate every four bars, creating a symmetrical eight-bar ‘block’ which repeats four times.

This consistent symmetry is, on the one hand, a perfect structuring schematic for a song mode of listening, as it creates predictable units which we can anticipate, order and explain. On the other hand, there is no apparent ending or break to this symmetry until a full minute and forty

---

80 ‘One might take exception to this [the concept of groove and song modes of listening], arguing against the implied contradiction between them. Certainly, a perfect rock tune ought to score high on both counts’ (Danielsen, 2006, p. 148).

81 Interestingly, after this eight-bar instrumental interlude, this perfect symmetry is relaxed in multiple ways. The order and division of lines between the rappers becomes varied. Warren G occasionally increases the rhyme density and introduces some divergent metrical structuring, and Nate Dogg extends his sung delivery by using longer lines spanning two bars instead of one.
seconds have passed, meaning that the unrelenting repetition of each structural unit at every level of periodicity is likewise pushing towards a groove mode of listening. Even though the listener becomes primed to expect each recurring rhyming line pair, every switch between the rappers, and the perfect symmetrical organisation of the bars, there is no alternative formal unit which shows up to demarcate any symmetrical grouping as a distinct ‘verse’.

Repeated symmetrical periodicities like those in ‘Regulate’ clearly help structure the listening experience. Whether this type of experience is one of ‘metricality’ or ‘form’ has been discussed in many central works in the literature, including Lerdahl and Jackendoff’s *A Generative Theory of Tonal Music*:

> It may be objected that the listener measures and marks off a piece at all levels, and that metrical structure therefore exists at all levels of a piece. For example, the listener marks off a sonata movement into three parts; the time-spans created by these divisions form the piece’s basic proportions; the question is whether these divisions are metrical, that is, whether the listener senses a regular alternation of strong and weak beats at these levels. Does he really hear the downbeat beginning a recapitulation as metrically stronger than the downbeat beginning the development, but metrically weaker than the downbeat beginning the exposition? We argue that he does not, and that what he hears instead at these levels is grouping structure together with patterns of thematic parallelism, cadential structure, and harmonic prolongation. (Lerdahl & Jackendoff, 1983, p. 21)

The formal units of a sonata movement are both long and dissimilar from one another, which works against them being experienced as metrical in any way. How similar and how short must a unit be for a group of them to be experienced as metrical, then? Clarke separates ‘rhythm and meter’ from ‘form’ in terms of the way in which the latter extends outside ‘the perceptual present’, and when the exposition of a sonata movement begins, the beginning of the development is well and truly part of the perceptual past. However, in a track like ‘Regulate’, with its sheer volume of cyclicality and many different repeating levels, everything is tailored to extend a listener’s perceptual present for as long as possible. As Clarke writes, ‘metrical structures undoubtedly extend beyond the level of single bars’; however, ‘a comparatively low upper limit should be established (at around the 4 to 8 bar phrase level, depending on tempo, structural complexity, etc.)’, since ‘it is unrealistic to assume that listeners (or performers) continue to be aware of a regular and all-inclusive pattern of relative strength and weakness between events that are widely separated in the
music’ (Clarke, 1987, p. 214). When ‘form’ takes over for ‘rhythm and meter’, the rhythms are no longer just felt or experienced in the present. Rather, they are perceived through a negotiation between the perception in the present and the (more or less) conscious recollection of preceding perceptual segments. London also points out that metres may contain ‘additional levels of entrainment, both higher and lower’ than their essential levels (the tactus, subdivisions of the tactus and grouping of beats into bars) (London, 2012, p. 17). In ‘Regulate’, it is not perfectly clear where this negotiation between perceptual present and past stops. One can probably entrain to the bar level, but how about to anything slower than that? Every four bars, when the loop loops and there is a switch between rappers, there is a clear structural ‘change’, but is this change *metrical*? Are the four-bar periodicities ‘hyperbeats’ in a ‘hypermetre’? I think not. The ‘hyperdownbeat’ of the beginning of Warren G’s sections is not a metrically stronger beat than beat 1 in Nate Dogg’s sections. Rather, these periodicities represent some other structural unit—they are symmetrical building blocks within a larger form but clearly part of ‘form’ and not ‘rhythm and meter’.

As can be seen in figure A6.1, the highest level of periodicity is the ‘Warren-then-Nate’ section, lasting eight bars. Since there is nothing in the music which demarcates the sixteen-bar period as a distinct section, there is no hint of a clear song structure. ‘Regulate’ teases our song mode of listening by giving us symmetry and predictability but makes it give way to the groove mode through sheer repetition. When the third eight-bar period appears (and is formally indistinguishable from the preceding two), it is as though the track urges us to abandon the expectation of further symmetry of form and embrace the being-in-the-G-Funk, enjoying the potentially endless back and forth between Warren G and Nate Dogg. ‘Regulate’ is uncommon in that its rhythmic constituents (the two-, four- and eight-bar periods) are exceedingly regular, but the formal units are not those we have come to expect of a hip-hop track. There is no chorus, and it does not really have what most rappers consider their standard unit of expression either—that is, a *verse*.

**Rap’s formal template: The verse**

While the concept of ‘verse’ has been repeatedly discussed throughout this thesis, it is, like ‘flow’ and ‘beat’, a term with multiple layers of meaning. In a rap context, ‘verse’ carries connotations beyond being a specific type of formal unit. For one, it has an elevated status in rap music compared to other genres, as it is the part of a track where a rapper really gets to express their craft. In addition to its importance, the archetypical rap verse has, over time, developed a strict formalised structure as well.
Although there are endless variations on a verse’s basic form, the majority of the rap songs recorded and released around the globe have verses that are sixteen bars long. There are many reasons for this, and the proportions of the various causality factors are worthy of dedicated research in themselves, but it is no surprise that rap verses trend towards a symmetric form akin to the groove-based music upon which rap is based (as seen in the example of ‘Regulate’). Whether it is the result of the use of samples or programmed loops, rap beats are typically cyclical in nature: two-, four- or eight-bar basic units are typically repeated, and, in many cases, verses and choruses are variations of the same basic loop.

Note, however, that the formalisation of a structure with verses of a set length is something which happened gradually as rap music became part of the musical mainstream. As mentioned in track A2, Condit-Schultz shows graphs that plot three features of rap song form over time: number of verses, length of verses in measures, and overall song length. (…) [T]here is an evident ‘standardization’ of rap form. Since 1995, nearly all rap songs on the Billboard Top 100 have had either two or three verses, with three being much more common. Similarly, around 1995 sixteen-measure verses became dominant. (Condit-Schultz, 2016c)

One can speculate as to why this standardisation happened. One explanation might be the shift in rap’s ‘function’ from live party and club music to recorded music for radio, CDs and MTV. The rapid growth of rap songs as mainstream pop in the twenty-first century might also have been a contributor to the formalisation of the form, as an enormous volume of instructional and (more and less) educational online content for aspiring (beat) producers and rappers has surfaced which often focuses on the sixteen-bar verse structure as a template. The processes of rap’s production might also contribute; rather than being ‘songwriters’, that is, rappers are often only responsible for the creation of the vocals on rap songs, with a much larger emphasis on ‘picking out the beats’ (made by beat-makers and/or producers) rather than making the beats themselves. In such cooperative endeavours, a clear, shared idea of the musical structure can ease communication. Lastly, the structures of global communication and commodification of the constituents of rap songs might influence the music in a profound manner as well. It is more and more common for so-called beat-makers (producers of beats

---

82 See Condit-Schultz (2016c, p. 144). As Condit-Schultz’s research shows, the trend might be changing, however, as the twelve-bar verse has become more and more prevalent throughout the 2010s and 2020s. Exactly how prevalent has yet to be empirically proven, and potential explanations for this shift will be discussed later in the track.
who do not necessarily do other types of engineering) to sell or lease beats online without much, if any, communication with the rapper(s). In such a situation, relative conformity to a common structure (not to mention stylistic or sonic features) might be necessary to connect with rappers with pre-written verses ‘looking for a beat’. Bottom line, the result of almost half a century of rap music is a common structure which is the foundation for most of the techniques, features, clichés and innovations of the rhythms of rap flows.

The prevalence-bordering-on-universality of the common structure is also present in the way rappers talk about flows. Edwards (2009, p 193) summarises what the large number of rappers he interviewed said about rap verses: ‘MCs often refer to the verses as the rap, the rhyme, the lyrics or “a 16”’. This also aligns with my own experiences. For example, one time when I was collaborating with another artist, he said, ‘Hit me up with a beat, and I’ll come back to you with sixteen’. The famous Rakim quote from the movie Something from Nothing: The Art of Rap (2012, 20:36) is a telling insight into a famous rapper’s writing process and his relationship to the common structure: ‘I try to start off with sixteen dots on a paper’. This type of conscious interaction with the boundaries and affordances of the common structure is clearly evident in the process of at least some rappers, and the negotiation between conforming to and diverging from the most common patterns within the common structure is an essential part of rap’s creative expression. As the sixteen-bar template became the standard, that is, rappers began to experiment more within the bounds of the rigid framework. Condit-Schultz notes that ‘there seems to be an increase in rhyme density until 2002 (...). The dramatic increase in rhyme usage in the early 1990s may be associated with the transition between old-school and new-school rap’ (2016c, p. 136). He also uncovers a significant variability among emcees when it comes to ‘phrase length’ and rhyme entropy (or the distribution of rhyme position within the metre; see track A2), wherein the trend is that artists showing the most variation tend to have released their music after the standardisation of the common structure of the verse.

This increase in rhyme density and entropy indicates that there is room for structural variation within this sixteen-bar framework. Nevertheless, a symmetrical organisation of structural building blocks remains the most common alternative by far. The convention of the line/bar coincidence and the prevalence of convergent metrical structure indicates a symmetrical organisation at the simplest multi-bar level (groups of two bars), and when this symmetry is challenged through the use of divergent metrical structure, the higher levels of symmetrical groupings of bars become the anchor points for the flow’s structure. If the two-bar unit
boundary is not adhered to, then the four-bar unit boundary typically will be. If the four-bar unit does not coincide with a line ending, then the halfway point of the verse—the eight-bar unit—probably will. I label this tendency of the flow to adhere to the larger symmetrical units the gravity of the form. This adherence also depends on the musical background (the ‘beat’), of course. Simply saying that the flow always operates in conjunction with a symmetrically consistent and predictable beat would be an oversimplification—as shown by, for instance, Ohriner (2016a), as discussed in track A2—but in general the beat emphasises the common structure of sixteen bars divided into two groups of eight and four groups of four, and it is the flow’s prerogative to destabilise this symmetry for aesthetic effect. I will return to these concepts on side B, as track B1 focuses on techniques which play with the smaller building blocks (bars, two-bar units and the boundaries between such units), while track B2 looks at various interactions between flow and the gravity of the form and how the former can both emphasise and contradict the latter.

Rap’s multiple (and flow’s own) musical metres

The gentle tug-of-war between flow and beat, song and groove, stability and instability, symmetry and structural tension is certainly a feature of the larger structural units of verses and symmetrical blocks in rap. Not so immediately obvious but still influential are the rap flow’s idiosyncrasies at the level of musical metre and subdivision. While we typically attribute organising principles and reference structures like this to the entire composite auditory stream, there is insight to be gleaned by looking at various layers in isolation as well as when they interact. Like a given rap’s communication of song and groove modes of listening, the beat and the flow often push in slightly different directions.

Almost all rap beats are in ‘an unrelenting duple metre’ (Ohriner, 2016a, p. 155). Of course, given the sheer amount of recorded rap music, exceptions are certainly present (in this thesis, an example can be found in track B4), but by and large, rap is in 4/4. But what does this ‘4/4’ really entail? Like other ‘rhythmic’ genres, the 4/4 metre of rap music has a particular balance to it. There are four heavy beats which have different qualities, archetypically expressed as the alternation of kick and snare drum which is onomatopoeically referred to as boom-bap in hip-hop discourse. This means that it is the off-beat positions between the four beats which represent the ‘weak’ beat positions capable of creating syncopation, cross-rhythms and counter-rhythms. In addition, as was the case for the funk music preceding it, there is a strong gravitational force towards ‘the one’—the first beat of the bar—and particularly the first downbeat of a basic unit. Additionally, hip-hop and rap, like funk and many other styles,
typically articulate a fast subdivision of the beat into four parts or more (that is, sixteenth notes or faster) in at least some elements of the music. Following J. H. Kwabena Nketia and his *The Music of Africa*, I call this subdivision level (the fastest one articulated) the *density referent* (Nketia, 1974, p. 127) All of this has, as I will show, ramifications for our metrical experience.

London has developed a theory about metrical competency he dubs ‘The Many Meters Hypothesis’:

> A listener’s metric competence resides in her or his knowledge of a very large number of context-specific metrical timing patterns. The number and degree of individuation among these patterns increases with age, training, and degree of musical enculturation. (London, 2012, p. 182)

In short, we have a quite sophisticated metrical repertoire which allows us to differentiate between many different types of metres, even when those metres share certain fundamental qualities. There is already a host of different 4/4 metres, including the vast range of rock and rock-and-roll and boogie variants, the different types of swung jazz styles, a range of folk dances from all over the planet, the four-on-the-floor electronic dance music types and, of course, the branch containing the soul, funk, disco and hip-hop metres most relevant to this thesis. The basic premise is that a trained listener can differentiate between metres (as opposed to holistic musical expressions) both quickly and subconsciously. For the different kinds of rap musics, which are often based on sampling or adapting different musical genres, this raises some interesting questions regarding the relationship between beat and flow.

Condit-Schultz shows that rap’s average accompaniment tempo (that is, the tempo of the beats) has steadily slowed throughout rap’s history (note that Condit-Schultz’s corpus ends in March 2015). While there is significant variation, the average seems to have declined from somewhere north of 100 bpm to somewhere between 80 and 90 (Condit-Schultz, 2016c, p. 135), with the outliers heavily skewing towards higher tempi (that is, there are significantly fewer tracks showing a lower tempo than the ‘average’). Condit-Schultz chooses the boom-bap/backbeat as the decider for the tempo, but he acknowledges the potential issues regarding

---

83 Condit-Schultz used a Bayesian approach where ‘a multi-level linear regression model was estimated using a Markov Chain Monte Carlo (MCMC) method’ (Condit-Schultz, 2016c, p. 135) rather than a simple average calculation, as the latter was clearly unsuitable.
double- or half-time ‘feels’ as well as the fact that the flow and the beat might operate with different metrical considerations:

Tactus interpretation has important repercussions in later analyses, especially when comparing songs with very different tempos. Prototypically, rap flow is organized around phrases, repeated patterns, and rhymes every measure. However, at faster tempos (especially 130 bpm or higher) flow may be structured principally around two-measure units—*up-tempo* flow. Conversely, at slower tempos (especially below 72 bpm) flow may primarily be structured around two-beat phrases—*down-tempo flow.* This suggests that songs evincing up-tempo flow could be notated in diminution (with the tempo halved) and vice versa for down-tempo pieces, putting all pieces in a ‘one-measure, one-phrase’ framework. (Condit-Schultz, 2016b, pp. 30-32)

This anticipates some current considerations regarding the *dual-planed* and/or ambiguous character of some rap or hip-hop metres as a key part of their aesthetic. There are several different factors determining which periodicity in a given metre will appear to be the primary one—the tactus—and in some styles, tracks, contexts or listeners, two different tactuses are viable, making them interchangeable or coexistent.

Consider the example from the track titled ‘a lot’ which was used in figure A4.2, for example. Here (reiterated as figure A6.2), it is in ‘half time’ relative to the transcription from track A4.

![Figure A6.2: Transcription of bars 1–2 in J Cole’s verse on 21 Savage’s “a lot” (from 2:37).](image)

The boom-bap tempo of the track is 73 beats per minute and the double-time tempo is 146.

Neither of these tempi is particularly close to what London identifies as the point of ‘maximal pulse salience’ of 100 bpm (2012, p. 31). While Western listeners display a preference for ‘periodicities in the 500–700 [80–120 bpm] range’, we are capable of perceiving and

---

84 Note that other scholars operate with other ‘ideal tempi’; de Clercq (2016), for example, argues for 120 bpm.
85 This may well be generalisable to all humans, but the experimental research that is the basis of this claim has in general been performed with Western subjects.
entraining to pulses ‘within a range from about 250–300 ms (240–200 bpm) to 2 seconds (30 bpm)’ (London, 2012, pp. 30-31). Thus, both the slow boom-bap tempo and the fast double-time tempo of ‘a lot’ are perfectly viable for a rhythmic piece of music, and neither is particularly close to the preferred range—that is, potentially likely to overpower the other. Instead, other factors decide which tempo is the tactus, as Condit-Schultz alludes to above. The lineation and rhyme placement, for example, advocate for the slow tempo transcription, as the rhymes end up on the expected beat 4 position, and no lines span multiple bars (as is the case in the 146 bpm version in track A4). The density referent of sixteenth-note triplets (or sextuplets) is also a better fit for the genre than the fast version’s eighth-note triplets. Why, then, choose the fast tempo notation in chapter A4? There are certainly musical arguments for it, such as the fact that the double-time tempo of 146 is reinforced by a steady, dynamically even hi-hat pulse which is easier to ‘feel’ as eighth notes than as sixteenth notes. However, the biggest reason is linguistic. In track A4, the excerpt from ‘a lot’ is used to compare prosodic subdivision to musical subdivision, and the most successful visualisation of this relationship involves the beat level wherein musical beats and poetic feet best correspond. Here, that is, when we want to foreground the musical metrical framework, the tactus of the excerpt demands reconsideration. In addition, the reuse of the same musical example with two different notations makes two separate points: (1) a certain type of rap metre (let us call it ‘trap’, even if that is an oversimplification) is dual planed, but the dominant tempo will be the slower one; and (2) using traditional notation in rap analysis is useful but has certain implications.

A specific type, or perhaps rather a branch, of hip-hop metre is rooted in the modern trap style popularised in Atlanta throughout the first decade of the twenty-first century. Grossly simplified, it is characterised by a very low boom-bap tempo which one could well consider ‘half time’, booming tuned and elongated 808 kick-drum sounds (or simply bass synth sounds reminiscent of the commonly used manipulated Roland 808 sounds) and very busy hi-hat patterns (that is, patterns with a high density referent). Mostly but certainly not always, the rap flow favours a triple subdivision (or the ‘triplet flow’ type discussed in track A1). While not all ‘trap music’ displays the signature characteristic of a dual-planed metrical framework with a slow dominant tempo, the archetypical trap beat does, as Norwegian hip-hop producers Kvilam and Kholebeatz emphasise in my interviews with them.\(^{86}\) Kvilam notes, ‘Of course, I

\(^{86}\) Erlend ‘Kvilam’ Lyngstad is one of the most prominent Norwegian hip-hop producers in the Norwegian scene at the time of writing. Kristian ‘Kholebeatz’ Hole is another incredibly prolific Norwegian producer with a more international profile, having collaborated with a wide array of rappers from the American South as well as the
work in double time, so if I make a beat that’s 110 bpm, it is slow in a way, right?’

Kholebeatz compares trap and electronic dance music (EDM):

But then there’s also a connection between trap and EDM with that, in my head, there’s not so much space to be behind [the beat], and that in trap and in EDM there’s almost the same bpm—only that trap is divided in half. If the snare in EDM (…) is on [beats] 2 and 4, then you can just divide it by two, and it becomes a trap-rhythm. So those genres go hand-in-hand, really. For instance, what Lil Jon does now is—it’s almost like crunk on EDM beats. So, there are incredible likenesses here. (…) It’s like basically the same, but (…) just by changing the snare pattern, you have a different genre.

The salience of the slower tempo as the dominant one also depends on how slow it is—the 55 bpm of Kvam’s hypothetical example is clearly much lower than the ‘preferred range’ for beat entrainment in a musical vacuum, whereas the double-time interpretation of 110 bpm is close to ‘maximum beat salience’. This may explain why many listeners will entrain to the double time even when other metrical considerations contradict its tactus status. Comparing this situation to lineation and lineation evidence, we see that trap metres are characterised by the presence of an alternative tempo to the dominant one, and there is an increased ambiguity between the tempi as the slow dominant one becomes slower and other typical characteristics (like the placement of snare hits or rhymes) undergo some deviation. There is some question as to whether one has entered the realm of another type of hip-hop metre altogether when this ambiguity disappears. In relation to contemporary trends in hip-hop production in 2018, the producer group Basmo Fam\(^7\) noted that the tempi in their productions ‘have gone from maybe 60, 70, 75 [bpm] to over 80. 85 and 90’. With the latter bpm values, there is no real possibility for a true dual-planed time feel, as the double time would be too fast and the half time would be too slow.

One consequence of rap music’s density referent being quadruple or faster is that the ‘weak’ metrical positions of the duple subdivision—the ‘off-beat’ eighth notes—become relatively strong in the flow. Rhythmic phrasings with syncopations, cross-rhythms and counterrhythms

---

\(^{7}\) Swedish superstar Petter. The author interviewed both rappers in September 2018. Interviews are transcribed by Eirik Jacobsen and translated by the author.

\(^7\) Brothers Øyvind and Amund and cousin Eirik constitute Basmo Fam, which has produced many tracks for a host of Norwegian rap acts, particularly those affiliated with sibling labels Oslo Records (of which Øyvind is CEO) and YGMG. This interview was conducted by Anne Danielsen and the author on 17 September 2018, transcribed by Eirik Jacobsen, and translated by the author.
will tend to use the weak sixteenth-note positions (or weak sixteenth-note triplet/sextuplet positions) because the eighth-note subdivisional layer does not communicate an *off-beat feel* unless the density referent is temporarily lowered. In a sense, one could think of the flow in isolation as a sort of 8/8 metre, as there are eight relatively strong metrical positions here. Of course, this does not imply that rap metres are not in 4/4, but rather that there are four distinct periodicities which define the metrical framework(s)—the bars, the beats, the duple subdivision of the beats and the next level of subdivision (the density referent)—which can be duple or triple.\(^88\) This superimposition of 4/4 and (a theoretical) 8/8 represents another way of conceptualising rap’s propensity for dual-planed metres—even if the dominant tempo (or, perhaps, ‘dominant tactus’) is clear in most cases.

**Transcription, traditional notation and quantised representation**

The use of traditional Western music notation for transcribing rap flows has long been a topic of discussion in the field of rap analysis. Some scholars, like Kautny and Komaniecki, use it as their preferred method of visually representing rap; others are quite dismissive of the practice, for several reasons. Ohriner, for example, believes traditional notation to be unnecessarily complex:

> While the rhythmic features of Western notation might be useful in transcribing rap music, the resulting transcription has far more dots, lines, beams, and flags than would seem necessary to convey rhythmic structure. And while I do not address the pitch content of rap flows here, Western notation’s approach to pitch would seem unhelpful as well. (Ohriner, 2019b, p. 8)

Rather than using traditional notation, then, Ohriner prefers his ‘molecular method’ (as Page calls it; see track A2), which is quite similar to the various grid systems which divide each bar into sixteen parts.

Adams’s scepticism towards traditional notation is not rooted in its complexity, ironically, but rather its lack thereof. In his chapter on rap analysis in *The Cambridge Companion to Hip-Hop*, Adams dismisses traditional notation due to the fact that it does not represent microrhythm in a satisfactory manner.

---

\(^88\) In some cases, where the dominant boom-bap tempo is slow enough, the density referent might even be quadruple, implying a fifth level of periodicity to entrain to. An example is Norwegian rapper Lars Vaular’s ‘Sjefen e tebake på jobb’ (2009), analysed and transcribed in Oddekalv (2017, pp. 67-72 and ‘vedlegg 9’). Kendrick Lamar’s ‘F*ck Your Ethnicity’ (2011), analysed in track B2 (page 175), also employs this thirty-second-note subdivisional layer.
The analyst must therefore make certain choices: in the rhythmic domain, he/she must either transcribe the rhythms faithfully, leading to a dense stream of sixty-fourth and even one-hundred-twenty-eighth notes (and tuplel variations thereof), or decide to quantize the rhythm to some degree, thereby gaining clarity of notation but sacrificing the rhythmic intricacy that energizes so much of the music. (Adams, 2015, p. 122)

This view is not really compatible with the understanding of rhythm which applies in this thesis. Categorical rhythm perception proposes that we do not experience the ‘rhythmic intricacy’ of microtiming as 64th or 128th notes but rather as modulations of quantised rhythmic categories. While it is true that quantised representation tools like traditional notation, sixteenth-note grids (which Adams himself employs) and the like do not communicate microtiming, but there are other features of the rhythmic structure that they do communicate, and very well.

This is at the core of this thesis’s division of representation into zoom levels. Rather than attempt to find or create ‘a universal analytical method for hip-hop music’ (Adams, 2015, p. 121), rap analysts have taken a different approach. Adams professes, ‘Analytical approaches to hip-hop have (. . .) been developed in the same ad hoc way that approaches to Western art music were: an analyst wishes to illuminate some aspect of the music, and fashions an analytical apparatus that will allow him/her to do so’ (2015, p. 121). In this regard, I believe the resulting analytical output is much richer than if analysts had formalised a shared (and therefore prescriptive) analytical framework sometime earlier in the field’s development.

The suitability of traditional notation for rap analysis does not depend on whether it is appropriate for analysis at the macrorhythmic and microrhythmic zoom levels, because that is not what it is for, and we (should) have other tools for that. Instead, the analyst must ask whether traditional notation contributes something to whichever quantised zoom level–type analysis one is doing, or whether one would be better off using a grid system or something else. Another misgiving regarding traditional notation is that its use could possibly reinforce a White racial framing for the analysis, since the system was not developed for analysing rap but rather for analysing score-based music from a diverse array of European traditions. To me, though, this is a bit of a dead-end argument, unless it is brought to bear against quantised representation as a whole. If the problem is that traditional notation overemphasises certain rhythmic features while ignoring others (which is valid enough), then the same criticism can and should be directed towards grid-based systems as well. Yet another misgiving about traditional notation is that it requires a certain level of musical training, which makes it less
accessible to a potentially interested untrained audience. Again, I am not so sure—music notation reduced to its rhythmic constituents alone is not that much more complicated than the alternative quantised systems. While the different types of subdivisions are indeed represented by stems rather than graphic spacing, traditional notation is much more flexible in terms of mixing types of subdivision, for instance.

My response to all these questions is to choose different tools for different tasks. When working at the quantised zoom level, I favour traditional notation when looking at rhythmic motifs, syncopation, off-beat phrasing and so on—that is, features internal to the flow. I reserve the use of grid-based systems for instances where the surface structure of the flow is lined up against something—perhaps the flow is lined up against features of the musical background, for example, or different symmetrical sections of the flow are lined up against one another. As should be clear from track A2, rap analysis generally spans all the various zoom levels, and there are very few instances of traditional notation being applied in unsuitable ways. Still, acknowledging the fact that much rap analysis has been focused on the quantised zoom level, the analyses in this thesis apply the wider and narrower lenses of the macrorhythmic and microrhythmic zoom levels to a greater degree than the quantised one.89

**Prominence, stress, accents: A topography of rhythmic events**

We will not, however, let go of the quantised zoom level and its rhythmic constituents just yet. Categorical rhythmic units interact with their surrounding frameworks, whether metrical, linguistic or otherwise, to create the rhythms of the music we analyse. In rap flows, these rhythmic units are syllables, and the concept of *prominence* means that some of these syllables tend to stand out in some way. In linguistic and poetic analysis, the prominent syllables are the ones which are *stressed*, whereas musical analysis tends to focus on *accents*. In the analysis of rap flows, the two will often converge—an accented note (a rhythmic unit) is nearly always a stressed syllable (a linguistic unit)—but not all stressed syllables are accented. Additionally, there are other modulating factors involved in how prominent a rhythmic unit actually is. Together, these factors contribute to a *topography of rhythmic events*, wherein different degrees of prominence—a combination of different types of accents—shape the rhythmic landscape of the rap flow.

---

89 If a reader is particularly interested in the quantised zoom level, my master’s thesis (Oddekalv, 2017) has that lens affixed throughout most of it. Similarly, Komaniecki’s doctoral thesis (2017) is a thorough study of quantised features of rap flows.
Broadly speaking, there are two categories of accents: those derived from the structural framework within which a rhythmic event exists, and those which are the result of some quality of the rhythmic event itself. Lerdahl and Jackendoff differentiate between two kinds of the former category—structural accents (those caused by the melodic/harmonic points of gravity in a phrase or section and especially by the cadence, the goal of tonal motion) and metrical accents (any beats which are relatively strong in their metrical context)—and group all of the latter under the umbrella term phenomenal accents, or ‘any event at the musical surface that gives emphasis or stress to a moment in the musical flow’ (Lerdahl & Jackendoff, 1983, p. 17). With this definition, of course, any kind of linguistic stress can be considered a type of phenomenal accent.

In linguistics, the relative stress of syllables is formalised into multiple distinct levels. For lexical stress (the relative weight within a single word), syllables defer to stressed feet (like in poetry), and ‘the foot of the feet’ (that is, the stressed event at the next hierarchical level) is the prosodic word or ω-word (see figure A6.3).

![Figure A6.3: Linguistic prosody. Hierarchy of lexical stress.](image)

These stresses might be realised in different ways including, most commonly, increased loudness, change in pitch, increased vowel length or some combination of these. They are analogous to the different flavours of accents used in music studies: the dynamic accent from an increase in relative volume, the tonic accent from a rising pitch, and the agogic accent from an increase in duration. The relative topography in linguistics is even more complex than this, in that there is also sentence stress (or prosodic stress, exposing the overlap between the disciplines of poetry and linguistics), and the analysis of the hierarchy of stresses is often extended to both a sentence and an even longer ‘utterance’ level. In music theory, there is rarely this level of nuance to the hierarchy (or topography) of rhythmic events, even if there are nuances to articulation marks in traditional notation (accent marks, marcato, tenuto and so on). The different types of articulation are rarely marks of different levels of prominence; instead, they represent different realisations of a single binary feature as either stressed or unstressed (or they mark articulation which does not necessarily have anything to do with prominence per se). As is the case with the analysis of poetic metre (a traditional two-level
system of *scansion*), the presence or absence of ‘accent’ only tells whether a note or syllable is stressed or unstressed. There are also several types of scansion which use three or four levels of stress—intended to analyse both poetic metre and poetic rhythm—but these are not particularly applicable to the analysis of rap flows. For one thing, scansion is generally an interpretation of the potential realisations of poetic rhythm, and there can (and, many would argue, *should*) be multiple alternative scansion of a single line of poetry. In a recorded or performed rap flow, the rhythm has already been realised by the rapper. Also, as discussed in track A4, a rap flow is typically structured according to a musical rhythmic framework, meaning that the musical metre, not the prosody, dictates the metrical framework. Instead, rap’s rhythmic expressivity revolves to a large extent around the ways in which the prosody interacts with the musical rhythmic framework, where it is free to place its stresses virtually anywhere. Rap’s musical dimension(s) will tend to level out or (at the very least) compress the distance between different levels of prosodic stress, which turns a multilevel stress analysis into a quite challenging epistemological prospect (which I will exemplify shortly and also discuss further in track B5).

In the rap analysis literature, there are slight variations in the annotation of accents. Beyond the simple indication of whether a syllable is accented or not, some analysts differentiate between different types of accents in their notation. Regarding the many types of grid-based visualisations, if there is any representation of prominence, all instances of verbal stress will typically be annotated. Among the analysts presented in track A2, Adams, Oliver and Ohriner use this approach, with slight variations. Oliver shows ‘Rakim’s emphasis’ rather than all verbally stressed syllables, and Ohriner distinguishes between the accents derived from computer analysis and those which are manually annotated. Condit-Schultz (2016c, pp. 129-130) applies analytical tools from the field of linguistics and distinguishes between pitch stress and syllable stress (attributable to loudness, articulation and vowel length), as well as other ‘prominence effects’ from pitch (such as ‘pitch motifs’, or intonational parallelism, wherein multisyllable rhymes display similar pitch contours). These data are not represented visually, however, but are part of the online database of rap flow transcriptions that Condit-Schultz’s ‘MCFlow’ project has launched.

---

90 Note that this is a methodological separation, not an indication of a rhythmic, musical or experiential difference between the accents.

91 Condit-Schultz adapts the ToBI system, short for ‘Tone-and-Break Index’ (Beckman, Hirschberg, & Shattuck-Hufnagel, 2004), which enables the transcription of linguistic parameters like intonation and accent (stress/prominence) for the purposes of digital analysis.
Differentiation between pitch accent and other types of verbal stress is also present in some analysts’ transcriptions using traditional notation. As discussed in track A2, Kautny does distinguish among relative pitches in his single-line transcriptions of rapped passages, at times showing three or more levels of pitch. In my own transcriptions, I refrain from annotating pitch unless discrete musical pitches are present (that is, the rap is ‘sung’). This again begs the question of what the tool is most suitable for. Traditional notation excels at representing quantised rhythm and quantised pitch. For the analysis of relative and unquantised pitch (that is, pitch not amenable to standard categories), I prefer employing other tools. Kautny’s approach may still be perfectly sensible, however, depending on which musical features the analysis zooms in on.

Another prominence-inducing quality sometimes displayed by a syllable or groups of syllables is what Connor (2015) refers to as ‘poetic accent’—that is, rhyme. While Connor calls it an accent, though, it is really not like other types of accent. First of all, it is not a phenomenal accent in the sense of being local and able to affect a single syllable independently of its surroundings. As discussed in track A5, rhyme is always connected to another part of the unfolding rhythmic stream, as rhyme does not exist until there are at least two rhyme instances. As such, ‘poetic accent’ is in fact a hybrid accent of sorts—the result of both a quality of the rhythmic unit itself (it rhymes) and the larger framework of which it is a part (it has to rhyme with something). It can also affect a whole group of syllables (in the case of multisyllabic rhymes), making the entire group slightly more prominent. This characteristic is not found in other types of phenomenal accent, which increase the contrast between the accented event and its immediate surroundings. Thus, instead of a fully fledged accent-inducing quality, I prefer to think of rhyme (in this context) as a type of modifying or modulatory feature which acts upon an already present prominence. It cannot really create an accent by itself, as we do not experience unstressed syllables as rhyming if they are not part of a multisyllabic rhyme (wherein at least one syllable is inevitably already stressed). Nor does rhyme tend to be notated as accents in the literature; instead, it is typically indicated using some sort of addition to the visualisation. Many analysts employ colour coding of some sort; some use numbering or other identifying markers for different rhyme classes.

When these accents and other prominence-inducing features like rhyme are notated, it is clear that the topography of a rap flow contains peaks of varying altitudes, and I will exemplify this in figure A6.4, prefacing a passage from the analytical amuse-bouche towards the end of this track. While the interplay between verbal accent, rhyme and the positioning of the stressed
and unstressed syllables within the metrical framework are the foundations of the rhythmic techniques of the quantised zoom level, the nuances of prominence and the interaction of different types of prominence are challenging to both notate and analyse.

**Figure A6.4:** A topography of rhythmic events: bar 2 of Eminem’s ‘The Real Slim Shady’ (2000, from 0:20). In the illustration: bedrock represents the music metrical structure (of stronger and weaker metrical positions); the soil layer represents the syllables; the grassy knolls represent stressed syllables; trees indicate marked accents (that is, stressed syllables in weak metrical positions); the light drizzle of precipitation illustrates rhyme (or ‘poetic accent’). Additional analytical levels indicated below: blue question marks in ‘marked accent layer’ indicate accented eighth-note positions; ‘Level 2’ indicates a potential higher level of prosodic stress; ‘cross-rhythm’ shows a (common) repeated cross-rhythmic pattern which is not necessarily visible in the simple rhythmic notation.

The illustration of the topography of rhythmic events in just a single bar of music shows the complex epistemological considerations involved in the analysis and representation/transcription of prominence. The musical notation—the plotting of syllables onto their corresponding musical metrical positions—forms the rhythmic bedrock and soil layer, but it does not in itself show any phenomenal accents, only metrical accents. The notational feature which does show prominence inherent in the basic metrical framework is the accent marker above. The question, then, is when should one apply the accent marker? From what is this ‘marked accent’, as I have labelled it, derived? The prosody layer in figure A6.4 uses traditional scansion notation to indicate stressed and unstressed syllables, and the stressed syllables are represented by the grassy knolls in the illustration above. Given that
something marked (that is, something which stands out from its surroundings) is prominent, a stressed syllable in an unstressed metrical position—a phenomenal accent which does not coincide with a metrical accent—is clearly very prominent. Thus, I have chosen to mark stressed syllables in the weakest metrical positions—off-beat sixteenth-note positions—with an accent marker.

Representations like this one fulfil a central requirement of precise notation—to be clear, uncluttered and unambiguous. However, they demand quite significant tacit knowledge of the reader/viewer: (1) one must acknowledge the relative ‘strength’ of the metrical positions—the eighth-note positions in rap’s metrical framework with a sixteenth-note density referent are relatively strong; and (2) one does not notate stressed syllables if they are not placed in metrically weak positions. Following from this, one might wonder whether there is a second level of prosodic prominence which one should represent in some way. I have marked the syllables I believe might be experienced as ‘more stressed’ (or what a prosodist might call ‘primary stress’) with question marks, because I am hesitant to claim that this second level is particularly evident or that my experience is necessarily indicative of what others might hear.

Another consideration is whether one should notate a lower level of accent wherein stressed syllables coincide with the relatively strong but slightly weaker metrical position of the eighth notes. In figure A6.4 these ‘weaker accents’ would be placed on the blue question marks in the ‘marked accent’ layer, but I believe this would clutter the notation unnecessarily and possibly undermine the significance of the marked accents, which represent the most important feature to mark clearly in the notation.

One important rhythmic feature of this passage which is ‘hidden’ and must be derived from a reader/viewer’s existing knowledge of musical notation and stress patterning is the cross-rhythm notated at the bottom of the figure. The scansion itself does not show this, as there are two ‘open’ sixteenth-note positions where there is no rhythmic event (syllable). Similarly, unless one is mindful of the relative metrical strength of the eighth-note positions and the fact that they coincide with stressed syllables unless there is an accent marked either one position ahead or one position behind, this cross-rhythm will not be evident in the musical notation either. I believe that there is no great solution to notating these kinds of features in a standardised quantised notation system, so they should instead be annotated if they are significant to the analytical context. There is also the abiding question of whether patterns like this actually contribute to prominence (and if they do, to what degree) or are ‘only’ a product of other types of prominence.
This brings us to the final feature—rhymes. In the illustration, I have included rhyme as a light drizzle of precipitation to indicate that it is a modulatory feature of other types of prominence. Rhymes clearly stand out in some way, but they behave differently than prosodic stress and the markedness of prosodic stress in weak metrical positions. As previously stated, I find the colour coding of rhyme to be a good strategy for representation, as it clearly indicates that there is something modulating the levels of rhythmic prominence, but something different from stress which should not be marked with some kind of accent symbol. All in all, the nuances of rhythmic topography are not easy for an analyst to traverse, and it should not be considered a fault of systems of quantised representation that they cannot emphasise all details equally. I believe a single-line music notation with one level of accent and the use of colour to emphasise rhymes is a great basis for representation at the quantised zoom level—it merely requires both the analyst and the reader/viewer to reflect upon which rhythmic intricacies might not be explicitly visually evident.

**Microtiming, microrhythm and its measurement and rulers**

The aesthetic power of the intricacies of microrhythm is an important subfield of all studies of so-called groove-based music, as well as rap analysis—regarding both the flow and the musical background. In track A2, I presented approaches from the existing literature, and in track A3, I presented my own approach to the topic via the theories of Clarke and Kvifte. The following sections are intended simultaneously to clarify some of the theories and terminology of microrhythm research which are relevant for rap analysis, and to deliberately muddy the proverbial waters by troubling the boundaries between digital and analogue, expressive and structural, of expected convergence and of timing and sound.

The basic premise of microrhythm is that there is variation in the measured and/or perceived temporal placement of rhythmic events as compared to *something* (a reference structure, another ‘simultaneous’ rhythmic event, and so on). This variation is small enough not to change the categorical identity of the rhythmic unit itself (its digital quality), but there remains within-category variation (an analogue quality) which is both audible and of aesthetic/musical-semantic significance. However, there are some complicating factors to this condition. One of them is signalled by the difference between the similar and often interchangeably applied terms of ‘microtiming’ and ‘microrhythm’, as Guilherme Schmidt Câmara points out:
The suffix ‘-timing’ is heavily connoted with event onset locations or inter-onset intervals (IOIs), to the point that they are somewhat synonymous. However, the perceived timing of a rhythmic event does not always correspond to the onset of its physical signal and instead has been shown to be influenced by other sound parameters, such as intensity, duration, and frequency/pitch. (Câmara, 2021, p. 18)

There is a nuance, in that ‘timing’ to a great extent points to specific positions of measurable physical audio signals—that is, ‘onsets’ or the measured distance between onsets (‘IOIs’)—while ‘rhythm’ encompasses the interaction between the physical signal and the listener’s experience, including whichever virtual reference structures are used to make sense of the sound. As Câmara points out, there might well be a difference between the measured and the perceived temporal position of an event.92 This is the question of where in a sound does the sound happen.

In the significant body of research done into this phenomenon (see, for example, in chronological order, Vos and Rasch 1981, Gordon 1987, Danielsen et al. 2019) the central term, coined by Morton, Marcus, and Frankish (1976), is ‘perceptual centre’ or ‘P-centre’.93 In all the analyses and theoretical discussions surrounding microrhythm in this thesis, the temporal position of any rhythmic event (generally a syllable) is considered to be its P-centre. Given that the exact position and variability of a P-centre are difficult to measure, as it is dependent on a wide variety of factors, all the ‘measured’ P-centres are my approximations using a combination of the audio and various visualisations (waveforms and spectrograms) of that audio in software like Praat, Reaper and Melodyne. This method is clearly neither exact nor reproducible, but that is not necessarily epistemologically prudent either. The point is that measuring a signal is not measuring perception, and my approximations take into account what experimental research (like Danielsen et al. 2019) has uncovered about how different factors (primarily attack time and duration, but also pitch and spectral centroid) influence p-centre placement and variability. It is also prudent to point out that the musical effects which I attempt to uncover and explain are questions not of exact measurable deviations or disalignments but rather of trends, tendencies and feel. A similar approach to this lack of precision is taken by Chris Stover in his exploration of ‘beat spans’—flexible time spans

92 The exploration of how qualities of sound influence the perceived timing of a rhythmic event is the basis for the research group which supports this thesis—TIME: Timing and Sound in Musical Microrhythm, led by Anne Danielsen. For a deep dive into the topic, I would direct the interested reader to the group’s publications. At the time of writing, the perhaps most accessible summary of important findings and terminology is to be found in the first chapters of Câmara’s thesis (2021).

93 Note that Morton and colleagues’ original paper uses the American English spelling of ‘center’.
within which several rhythmic events might occur in the rhythmic weaves of Afro-diasporic music:

It is worth noting that there is no attempt here to construct a model that is in any way mathematically rigorous, since the nature of the music in question does not readily conform to, say, precise measurements in milliseconds, or fractions of distances along some sort of measured beat span segment. Another way of saying this is that the malleability of the musical flux, the subtle and nuanced way in which the two metric strata gently tug on the actual performed musical objects, is such that precise measurement is not only futile but also yields little useful information for the analyst. (Stover, 2009, p. 249)

Some theorists will find the preceding and subsequent narrative maddeningly imprecise, with such vague descriptive terms as ‘a little later than’, ‘just before’, and ‘shading toward’. But these are precisely the terms that are required for such descriptions. It is not the fact that an attack point is displaced by n% or by x milliseconds or by a y:z ratio as much as the observation that the gravitational force of the contrasting stratum is pulling the attack point away from a quantifiably measurable grid. (Stover, 2009, p. 250)

This embrace of inexactness does not mean that the microrhythmic zoom lens has a low resolution or that milliseconds do not matter. They do, and they will be measured. Rather, it is an acknowledgment of the imprecision and subjectivity of human cognition. Even though longer rise time and total duration seem to have the same effect upon listeners with different backgrounds and training (Danielsen et al., 2022), my P-centres and beat bins (a term I will shortly explain) are not the same as everyone else’s. 94 Thus, they are indicative of some sort of rhythmic experience which is fairly consistent within my own experience and likely evocative of the experience of others.

When analysing microrhythm, a recurring challenge is to consider what one is measuring and, perhaps even more importantly, which ruler one ought to apply when measuring it. Whether one is measuring ‘non-alignment’, as Ohriner (2019c) calls it, ‘discrepancies’, as conceptualised by Keil (1987), or ‘deviation’, as discerned by Gabrielsson (2003), there remains the important question of what it is that one aligns with, deviates from or measures

94 This can be seen in the variability of P-centre placements amongst experiment participants in Danielsen et al. (2019).
against in the first place. There are many different strategies informing this enterprise, depending on the types of music or performance being studied, but generally one measures the internal relationship between rhythmic events and/or the alignment/deviation of rhythmic events from some sort of framework or ‘grid’. This ‘grid’ can either be an idealised one—typically fully isochronous—or one which is derived from other layers of a composite auditory stream (the distance, for example, between the syllables of a flow and the drum hits in the beat). Suffice it to say, there are epistemological challenges in determining what varies against what and whether there are mutual influences between two layers or a hierarchical relationship instead, not to mention where the exact points of measurement are and whether and how the relationship between layers influences the P-centres of the individual rhythmic events.

In the case of rap analysis, these rulers and idealised grids are easier to identify than in many other cases. Because rap music tends to be produced in a way that involves some sort of quantisation of the sound signal, whether through sampling, programming using quantised sequencers or ‘grid-snapping’ in a DAW, or even recording in layers over a metronome, at least some of its rhythmic spans will be perfectly regular. In the case of a sampled funk groove, perhaps the perfect quantisation point is at the loop point which is aligned with a basic unit. In a modern trap track, all the elements of the beat might be plotted in a perfectly quantised sequencer, resulting in complete mechanical precision to the millisecond at all density referent levels. In this thesis, each microrhythmic analysis will include a discussion of which ruler has been applied. The push-and-pull between different reference structures, between layers in the auditory stream, and between different perceptual categories are all central to the microrhythmic concepts presented. Thus, the inexactness of P-centre placement, rulers and categorical boundaries represents the epistemological ground upon which the analyses are founded.

Expected convergence, systematic variation and expression versus structure

Two similar pairs of terms which have regularly arisen in this thesis are digital/analogue and structure/expression. The former refers to the relationship between perceptual categories versus within-category variation and modulation. In the latter, ‘structure’ is used to describe rhythmic categories and thus implies some kind of cognitive and/or representational quantisation. Not all analogue qualities are ‘expressive’, and depending on how one defines different rhythmic categories, the analogue might well be structural. ‘Expression’ (as I use it here) implies at the very least an exaggeration or suspension of structure. It might take a
structural quality and expand upon it, but it is not consistent or repeated enough to become structure itself. This is how Ingmar Bengtsson, Alf Gabrielsson and colleagues explain it in their ‘Empirisk rytmforskning . . .’ (1969). Of the various ‘SYVAR-D’—systematic variations in duration—some are tied to specific expressive underpinnings of structural features of the music (that is, they are ‘expressive’ variations), while others are part of the music’s rhythmic ‘dialect’. Among the many examples of such structural systematic variations in duration, the one most applicable to the study of rap is swing.

Swing is a huge category of more or less consistent variation in rhythmic duration between alternating rhythmic units of one subdivisional category. The archetypical jazz swing is on the eighth notes, where the ‘on-’ eighth notes are longer than the ‘off-’ eighth notes. A full blues ‘shuffle’ is approaching a quantised 2:1-ratio, where the ‘on’ is twice as long as the ‘off’. In funk music, the sixteenth notes are typically swung, and this tendency carries over to rap flows, where the sixteenth-note-level density referent is naturally the level of durational variation. There are also many degrees of swing, whether mechanically programmed or more or less inconsistently performed. Some hip-hop grooves use a ‘quintuplet swing’ with a ratio of 3:2 or even septuplet (4:3) or other programmed ratios at either eighth- or sixteenth-note levels. This variation occurs most often in the hi-hat or equivalent (various shaker-type instruments, for instance) timekeeping instrument layer, and it may or may not be consistent between instrument layers.

In short, swing is a structural rhythmic phenomenon which depends on analogue qualities within a digital category. One way of thinking about it is via nested categories. ‘Eighth notes’ is a category. Modulated by the analogue quality of swing, a new category ‘swung eighth notes’ is created within the parent category. The sophistication (or persistence) of the given listener might uncover more specific subcategories such as ‘inconsistently swung’ or ‘quintuplet swung’, and these categories can again be modulated by more or less systematic variation, whether dialectical or expressive. In the same way that London’s ‘many metres’ hypothesis is based on a listener’s metrical repertoire, the listener can have a repertoire of swing which is highly sophisticated. Combined with the fact that P-centre identification is both subjective and not strictly temporal, it becomes clear that the phenomenon of swing is not fully measurable from the audio signal, so it is imperative to stress the listener’s role in the formation of swing as a rhythmic phenomenon.

There are other consistent inconsistencies in microrhythm which one can find in the analysis of rap flows. In the interaction and expected convergence between events in the various layers.
of the composite auditory stream, aesthetic microrhythmic tension and ambiguity arise. One recurring concept in the analyses in this thesis is that of the beat bin. Rather than being precise points in time, that is, beats in fact have some sort of duration: ‘A beat bin can be understood as the temporal interval surrounding every metrical beat (and metrical subdivision) where a sonic event would be perceived as being played “on” the corresponding metrical event’ (Haugen & Danielsen, 2020, p. 357). In other words, multiple rhythmic events occurring within a certain temporal distance from one another can all belong to the same ‘metrical event’—they are all ‘on-beat’, even if their P-centres do not align. Some musics have very narrow beat bins (that is, a rhythmic event can be very close to the temporal position of ‘the beat’ and be experienced as categorically ‘not part of the beat’), while others have wider beat bins. The latter is typically a product of some sort of categorical ambiguity within the individual layers of the composite auditory stream which influences the experience of the whole, as well as extensive repetition which establishes the dominant interpretation of the part-of-the-beat and not-part-of-the-beat categories. Clear examples of how this occurs in (groovy) musical contexts are found in Danielsen (2010b), where the term ‘beat bin’ originated in an analysis of D’Angelo’s ‘Left And Right’ (2000), and in (Danielsen, 2018), where another D’Angelo track, ‘1000 Deaths’ (2014), is the primary example.

For this thesis, the concept of the beat bin is particularly interesting as a local rather than a recurring phenomenon. This is because the layer of the composite auditory stream which is the main analytical focus—the rap flow—is in general more ‘expressive’ and less structurally consistent in its microrhythmic variation than the instrument streams of Danielsen’s analyses. Rather than the consistent reinforcement of repeated basic units, we find that wide beat bins in the interaction between rap and musical background arise in relation to the expected convergence of syllables’ metrical positions and the musical background’s evocation of the corresponding metrical framework. In track B3, the analyses will show how syllables can communicate categorical belonging to a metrical position within the framework of the flow as it unfolds in time, and how there can be friction between various categorical reference structures, creating effects like wide beat bins or other aesthetic ambiguities.

Microrhythm tugging at the edges of perceptual categories raises certain epistemological questions. Since music unfolds in time, our reference structures, experiences, expectations and so forth are constantly shifting because they are informed by what has just happened, what happened a while ago, and how what happens next fits or does not fit with these shifting frameworks. Even if the physical signal is identical (a perfect programmed copy or sample),
the experience of a beat’s bin will inevitably be different in the first bar of a track than it is in the fifty-sixth. And what about the hundredth time one listens to the (whole) track? Is the experience the same as the first, tenth or four hundredth time? Is the experience significantly influenced by the listener’s mode of listening? Will the avid fan have the same or similar experience of a microrhythmic phenomenon as the analyst?

This conundrum also rears its head in the studio, both when assessing different vocal takes and when programming/editing beats. Repeated re-listenings and pointed attentional focus will occasionally shift the experience of whether an event is ‘part of the beat’ or not. Here, different perspectives can play a huge role. In one specific case, I kept nudging a shaker sound further and further ‘behind’ the grid—to the point where I felt that the rhythmic friction between the shaker and the kick and snare made the shaker start to feel like ‘not part of the beat’. My bandmate Bennærn, however, was more ‘detached’ from the temporal minutiae in his listening and exclaimed: ‘That’s dope! You can take it even further behind’. While I was initially befuddled by this—as it contradicted my own feeling (which was rooted in the close listening and active music-making mode I was in)—acting on his impulse and moving the shaker even further ‘behind the beat’ allowed me to shrug off my ‘not part of the beat’ interpretation and listen with fresh ears. It was indeed dope, and the shaker was clearly part of the beat again, albeit a quite wide beat.

The divide between flow and delivery and the questions of what is rhythmic structure and what is expression were all touched upon in track A1. At this point, fast approaching the end of side A, we find ourselves better equipped to explore these grey areas. Given that one can perform the same flow with different deliveries and the same rhythmic structure with different expressive variations, there seems to be a clear parallel between flow and structure, on the one side, and delivery and expression, on the other. This is one way to justify my omission of delivery in my definition of flow and to begin explaining why, for example, Edwards includes being ‘very slightly before or after the beat, rather than exactly on the beat’ (2009, p. 256) as part of delivery but not flow. Similarly, Adams (2008) considers ‘the extent to which the onset of any syllable is earlier or later than the beat’ as an ‘articulative’ rather than a ‘metrical’ feature of rap in his flow definition. Clearly, expressive timing seems to overlap more with delivery than with flow in these scholars’ work.

However, as I keep coming back to, there is a floating boundary between what is ‘structural’ and what is ‘expressive’, and our experience of where a rhythmic event is (its P-centre) relies
on both temporal placement and other features of sound. Consider the following quote from Norwegian rapper Oral Bee:

We’ve always been on that ‘west coast wagon’ where people often have been laidback and been a little ‘bakpå’ [Norwegian word meaning ‘behind’] on the beat. Dragging out the words, especially at the end of the lines. (…) I don’t like when it’s toooooo much either, though, so the stuff is usually quite on-beat. (Oral Bee, 25.01.19)

First, the state of being ‘behind’ is presented as a consistent modifier for the rhythmic feel. It is, in other words, a systematic variation and thus arguably structural rather than expressive. Second, ‘laidbackness’ or ‘behindness’ seems to be primarily communicated by sound features other than the temporal placement of syllables. Oral Bee increases the duration of syllables by dragging them out, for example. This is consistent with findings in Câmar, Nymoen, Lartillot, and Danielsen (2020b) and Câmara, Nymoen, Lartillot, and Danielsen (2020a) which reveal that an increase in duration is one way musicians communicate laidbackness.

The waters are muddy, though. Since microrhythm sometimes communicates structure and other times expression, and delivery sometimes strongly influences rhythm, there can be no hard and fast rule in deciding what should be part of an analysis of a rap flow. In general, I will tend to analyse both rhythmic structure and rhythmic expression as well as the ambiguity of the boundaries between the two, depending on which feature is being analysed and which zoom level is suitable. In the analyses of flow, I will generally avoid extensive analysis of delivery, but when features which are typically part of delivery (like pitch, for instance) have strong rhythmic implications, they will be included. To round off side A, I will now present an analysis of a segment of a rap verse using the various zoom levels and theoretical frameworks presented thus far.

---

Oral Bee (Anders Kranmo Smedstad) is the foremost pioneer of westcoast ‘gangsta’ rap in Norway. As a rapper, he is amongst the most prolific recording artists in the country, releasing multiple albums every year, and as a producer (mainly working in tandem with fellow Da Playboy Foundation member Mr. Pimp-Lotion), he has collaborated with many artists from Los Angeles and the Bay Area in California. The interview was conducted and translated by the author.
Zoom levels in contrast and tandem: An analytical amuse-bouche

In this final section, I will apply different analytical lenses to the opening of the first verse of Eminem’s ‘The Real Slim Shady’, highlighting how it conforms to yet challenges the line-bar coincidence, how it uses rhythmic rhymes, and how it reinforces that parallelism using microrhythm and melody. I will start out with the macrorhythmic lens in figure A6.5, then zoom progressively further in.

Figure A6.5: Macrorhythmic representation of Eminem’s ‘The Real Slim Shady’, verse 1, bar/line 1–4, 0:17–0:26. Primary rhyme instances marked with red outline, weak alternative lineation from syntax marked with dotted light green. Lines’ distribution (red colouring) and rhyme positions within musical metre mapped on the right.

As the figure shows, the lineation structure creates a convergent metrical structure, with the poetic lines overlapping the musical metre in the coloured bar representation on the right side of the figure. Since there is no divergent metrical structure to speak of, the macrorhythmic representation in figure A6.5 (it omits information about most time value classes to emphasise larger rhythmic building blocks and features) is primarily useful for highlighting the structural influence of the positions of the primary rhyme instances (marked with red outline) and potential ambiguous lineation at the (linguistic) syntactical level (marked with dotted light green). While there is a convergent metrical structure and one single rhyme complex (there is only one primary rhyme class), there is a structural parallelism which groups each line pair as its own structural unit—namely, the repeated distribution of primary rhyme instances at the end of bars 1 and 3 and at both the beginning and end of bars 2 and 4. The lineation evidence is loud and clear, as it introduces these symmetrical structural units spanning two and four bars. The only feature which vaguely suggests another sort of grouping is the syntax, as lines 2 and 3 could well be interpreted as one long line. However, this is a very weak alternative, as all sorts of parallelism and genre conventions, as well as a pause inserted between ‘door’ and

---

96 Eminem is probably overrepresented in the rap analysis literature, even when considering his commercial success, longevity and cultural influence. The example here is chosen because it can efficiently showcase the application of a large number of analytical tools using a short segment of music.
‘and’, confirm the dominant lineation. While the metrical structure is convergent, the lengths of the lines are not equal either, due to pickups and slight variations in line endings. This can be seen in the colouration of the bar representation in figure A6.5.

While the macrorhythmic representation indicates a quite straightforward convergent metrical structure, the quantised representation of the same passage in figure A6.6 shows that the flow applies some gentle tugging upon the musical metrical boundaries.

For one thing, the primary rhymes, by virtue of consisting of a four-syllable compound rhyme class which spans more than a full quarter note beat, are not placed directly on the archetypical beat 4 position except for one time. The final rhyming syllable anticipates beat 4 in bars 1, 3 and 4, which means that the first rhyming syllable—the start of the rhyme class—is placed on beat 3. On the one occasion where the rhyme class begins ‘on the four’, at the end of bar 2, it is anticipated by a sixteenth note rather than placed straight on the beat. This rhyme position introduces a slight tendency towards a divergent metrical structure, as the final rhyming syllable (‘door’) could be heard as a syncopated one (a ‘one-rhyme’, a technique which will be discussed at length in track B1). However, due to the sixteenth-note density referent, the metrical position (final eighth note position in bar 2) is relatively strong, and the line does not quite feel like it stretches into bar 3. What this later placement of the rhyme does, though, is shorten the third line in comparison to the other three, which all have pickups. Since line 2 ends later in the bar, there is no room for a pickup to line 3. This shortening of line 3 is underpinned by the position of the rhyme which initiates line 4, which anticipates beat 1 (the first stressed rhyming syllable is placed on the last sixteenth-note position in bar 3). While the rhyme is still clearly ‘on the one’ and belonging to line and bar 4, it takes a

---

97 Note that in passages like this, where there is no real potential terminological ambiguity, I use ‘primary rhyme(s)’ or even ‘rhyme(s)’ to signify ‘primary rhyme instance(s)’.
chunk out of the already short line 3, emphasising the dissimilar line lengths within the convergent metrical structure.

The displaced rhymes of the four of bar 2 and the one of bar 4 highlight a couple of concepts which are key to this thesis, beginning with the recurrent theme of expected convergence. The stressed initial syllable in the primary rhyme instances, even augmented by ‘poetic accent’, brings with it a strong feeling of ‘on-the-beat-ness’, particularly if one follows the flow part of the composite auditory stream more closely than the beat and its constituents. If one attempts to rap the passage a cappella (or even listen to an a cappella version of Eminem’s recording) there is real ambiguity in terms of whether these stressed syllables before the bar lines are experienced as syncopations or the beats themselves are of uneven lengths and the stressed syllables dictate the positions of those beats. The latter possibility introduces all kinds of problems for any combined transcription of flow and musical background, as the beat positions would differ significantly, but an argument can be made that it would be a more faithful representation of some aspects of the musical experience. Thinking in beat-bin terms, there is an interesting ambiguity to these stressed rhymes’ status as either part-of-the-beat or not-part-of-the-beat because the latter interpretation depends upon the existence of (at least) two parallel musical metrical frames of reference—the flow and the whole (whether this is the case or not is very much listener dependent and evades any objective measuring). This points to how experienced beats are musical gestures in and of themselves: they do not necessarily correlate to measured temporal position or the expected placement within the representations used when discussing them. All of this emphasises the extent to which a quantised transcription is an interpretative act which always raises interesting epistemological and theoretical questions.

Which brings us to the second point: microrhythmic phrasing, ‘discrepancies’ and internal relationships among syllable durations can either clarify or blur these categorical distinctions. Whether conscious or accidental, the microrhythmic intricacies of a rapper’s flow will heavily influence the rhythmic experience with regard to both expression and structure.\(^{98}\) In the case of ‘The Real Slim Shady’, Eminem’s microrhythmic phrasing—exemplified in figure A6.7—is a source of tension and ambiguity, offering differing and contradictory structural interpretations and showcasing how loose the connection between the musical metrical and

---

\(^{98}\) And it is, of course, often unclear whether something is expressive or structural, and what consequences that will have for a quantised transcription. Thus, a really interesting experiment would be to compare different transcriptions of the same passage, discussing the reasoning behind the different choices (if any).
rhythmic frameworks of the flow and the musical background can be while still presenting a completely coherent and intuitive whole.

Figure A6.7: Eminem, ‘The Real Slim Shady’, verse 1, bar 2. Spectrogram and waveform graphic with lyrics mapped to interpreted syllable P-centres (found using a combination of Praat and Reaper), and blue lines showing deviation from the transcribed interpretation of figure A6.5 as plotted onto a mechanically quantised grid.

The microrhythmic representation of figure A6.7 shows just how much and how inconsistently Eminem’s timing deviates from a fully isochronous grid. Of course, such a grid does not accurately represent any actually sounding music or strict universal reference structure, so its exact values do not really ‘matter’. However, as in most rap, some parts of the musical background are mechanically quantised to certain points in such a grid, and the internal relationships between syllables are interesting to compare to the idealised average which a grid represents. One thing that leaps out from the representation in figure A6.7 is that there is no real internal consistency to the distance between events at the level of the density referent (the sixteenth-note level). There is no consistent swing ratio, and if we extrapolate the quantised transcription, the internal relationships within different rhythmic figures vary wildly from almost completely even sixteenths to a more than 4:1 swing ratio. This muddying of the categorical boundaries at the subdivisional level is part of what makes a quantised transcription such an act of interpretation rather than an objective truth in this case. When there is little consistent evidence to be found in the temporal placement and internal duration of rhythmic events, other considerations become more important when determining categorical belonging. For instance, both the beginning and the end of the line/bar in figure A6.7 are instances of the primary rhyme. And while the first, ‘all on the floor’, is interpreted as on-the-beat in the quantised transcription, the second, ‘burst in the door’, is interpreted as off-the-beat (anticipatory). However, looking closely, they appear to be placed at a similar distance from the quantised grid position of their respective beats (approximately a thirty-second note ahead). So, why do they have different categorical qualities in my interpretation of the rhythmic structure? The reasons range from parallelisms in the rhythmic phrasing to my sense of Eminem’s general microrhythmic tendencies and even to an epistemological
compromise. I hesitate to fully detach the analysis of the flow from the reference of the musical background (which might entail using longer and shorter beats and/or irregular time signatures, for instance), as the friction between the two is an important aesthetic element in the music.

Another fascinating feature of this passage and its inconsistencies at the sixteenth-note durational level is the apparent further level of the parallelism—that of *rhythmic rhyme*, as can be seen in figure A6.8.

![Image](image_url)

**Figure A6.8:** Primary rhyme gestures in the first four bars of ‘The Real Slim Shady’. All can be considered the same rhythmic figure (strong-weak-weak-strong). All except the sixth also converge towards a similar microrhythmic pattern (long-short-long-short/syncopated).

The figure shows that the microrhythmic pattern in five of the six instances of the primary rhyme figure are very similar and could be interpreted as ‘swung’. The difference between long and short in the third instance (‘burst in the door’) is even exaggerated enough that one might be tempted to revise the sixteenth-note notation as a dotted sixteenth note and a thirty-second note. The main reason for not interpreting the gesture as ‘something different’ from the other instances of the rhyme figure is that all the different interacting parallelisms invite a perceptual categorisation of all the figures as ‘the same’, particularly given that the swung microtiming pattern itself can be considered *marked*. Most of the repeated sixteenth-note patterns surrounding the primary rhyme figures are clearly more even, as can be seen in figure A6.9.
While rhyme in its basic form is speech-sound parallelism, and rhymes in rap are augmented by rhythmic parallelism, the layer of microrhythmic parallelism which Eminem adds to the rhyme gesture further enhances the markedness of the rhyme and the peaks of the topography of the rhythmic events.

In general, analysis at the microrhythmic zoom level will entail some sort of measurement of P-centres against some sort of ruler, either external (like the quantised grid applied here) or internal (measuring absolute distance between events). In the microrhythmic analyses on side B of this thesis, I tend to include one or more potential interpretations of quantised rhythmic figures in the various visualisations, rather than placing them in a separate figure, as is the case here. Finally, I present an example of a visualisation of a musical feature which is not really part of the flow or at home at any specific zoom level: the melodic content/pitch, with which one can inform the analysis of rhythm as well (see figure A6.10).
Figure A6.10: Melodic parallelism in Eminem’s ‘The Real Slim Shady’. Celemony’s Melodyne (version 4.2.4) software screengrab with added highlight and lyrics. Similar general pattern (high-low-high) attributing to the strong-weak-weak-strong prosody pattern.

While there is not an exact match between the two melodic gestures in a chromatically quantised sense, the melody of Eminem’s rap does not conform to a chromatic scale in any case. The melodic pattern of high-low-high on the rhymes stands out in the melodic context, stacking an additional layer of parallelism onto the rhyme gestures.\(^99\) One thing this showcases nicely is how the addition of rhythmic, microrhythmic and melodic parallelism reduces the need for any actual syllabic similarity between rhyme instances. The first syllable of the second rhyme instance (‘all on the floor’) is experienced as being part of the rhyme even though it does not rhyme phonetically with the corresponding syllable in any of the other rhyme complexes (‘all’ does not rhyme with ‘per-’, ‘burst’, ‘worse’ or ‘fur-’). As I said in track A5, there is more to rhyme than rhyme.

\(^{99}\) Note that there is some redundancy to the ‘stacking’ of parallelisms here. The way that high notes correlate with stressed syllables and low notes correlate with unstressed syllables is indicative of pitch being amongst the most important ways of realising stress. The melodic parallelism, then, is not truly a separate parameter from verbal stress in this passage.
SIDE B
Interlude

To use a topical metaphor: side A has been an (admittedly lengthy) set-up line, aiming to allow for side B to be as powerful and fulfilling a punchline as possible. The epistemological, methodological and theoretical discussions have (hopefully) set the stage for the analyses to be clear, nuanced and illuminating, without terminology and theoretical implications becoming cumbersome and confusing. This interlude will, with the utmost brevity, sum up the central points of side A, which inform the coming analyses.

- A rap flow is the rhythmic structure of the words and rhymes in the vocal track(s)/performance in a work of rap music.
- *Delivery* can and will influence rhythmic structure (and thus *flow*), but it is not in itself rhythmically structural.
- ‘Rhythm comprises an interaction between non-sounding reference structures (…) and sounding rhythmic events’ (Danielsen, 2010a, p. 4).
- The rhythmic experience is comprised of both rhythmic *structure* and rhythmic *expression*. This is related to (but not identical to; see discussions in tracks A1, A3 and A6) the interaction between *digital* and *analogue* features—the former denotes perceptual categories, while the latter denotes continuous within-category variation.
- Things which ‘stand out’ (that is, are *marked*) are dope.
- Musical analysis is intimately connected to musical practice (composition/performance), and any analysis is similarly connected to the analyst. My analyses are my takes on the analysed music—they are ‘covers’, if you will.
- The analysis and visual representation of rap flows can be broadly categorised into *zoom levels* which correspond to different levels of categorical rhythm perception.
- Music unfolding in time is a *composite auditory stream* with more or less ‘separable layers’, in the sense that the listener can attend more or less closely to them.
- Rap’s verses are typically symmetrically organised. They are most often sixteen (or twelve) bars long, and these bars are most often subdivided in a symmetrical manner into blocks of eight and four bars.
- *Expected convergence* between features of different layers of the composite auditory stream is a central structural mechanism for listening to rap flows.
- Musical metrical structure and poetic metrical structure are similar but not ‘the same’. Different levels correspond to one another, but they do not necessarily coincide. A
central aspect of rap flows is the way in which musical rhythm and metre and poetic rhythm (prosody) and poetic lines interact.

- **Lineation**—the division of text into lines—is an implied form. This means that lines are not absolute ‘facts of the text’ but rather experienced formal units (Fabb, 2002, p. 136).

- In rap flows, the most common relationship between bars and lines is that they coincide—I call this *convergent metrical structure*. Often, however, they do not coincide (*divergent metrical structure*), and this is the basis for many essential flow techniques.

- Rhyme is, fundamentally, a *speech-sound parallelism*, and in addition to being dope in and of itself, it is central to the rhythmic structure of rap flows—which is the aspect of rhyme which I analyse.

- Some rhymes are more structurally significant than others. These *primary rhymes* establish *rhyme complexes*—a type of formal structuring unit which interacts with metrical units (bars and lines) and larger groupings of these. Other, less structurally significant (but no less dope) rhymes are *secondary rhymes*.

Obviously, there are many other nuggets of wisdom in the nearly 130 pages of side A, and I will dutifully return and refer to central points throughout the analyses which follow. Still—with this brief summary fresh in mind—we should now be well equipped to tackle side B. The different tracks will present and take apart examples of different techniques and concepts rappers employ in their flows, and this will be done using the zoom levels and representational concepts we just applied to ‘The Real Slim Shady’ at the end of track A6. Note that the sampling of rappers/examples is made in a quite unsystematic manner—the examples are chosen because they showcase a technique or variation. I have tried to represent a broad variety of rap and rappers in terms of stylistic identity, gender, nationality and so on, but my background and preferences influenced the selection of examples, so there are clearly more examples from Norwegian rap and OutKast flows than one might find in the typical rap analysis-thesis.

Tracks B1 and B2 take place mainly at the macrorhythmic zoom level—focusing on lines, bars and larger blocks and sections of flows. Track B3 explores microrhythmic features, zooming in on within-category variation (‘analogue features’) and categorical ambiguity using a combination of events plotted onto grid rulers and other quantised representations to illustrate potential structural interpretations. The final analysis track, B4, looks at melody’s
rhythmic contributions, and track B5 sums up the thesis, addressing the research questions and the overarching mega-question: *What makes the shit dope?*

Note that representation affiliated with the quantised zoom level (that is, traditional notation) is used sparingly throughout these analyses to present different structural interpretations of variations in microrhythmic phrasing (in track B3) and to capture discrete musical pitches (in track B4). This is, as previously stated, a calculated response to much previous rap analysis, which has tended to favour this lens both graphically—in representation—but also analytically—that is, with regard to the features and techniques being addressed. Because the most novel methodological contributions of this thesis are concerned with the macrorhythmic and microrhythmic levels, the features and techniques which are best viewed through those lenses are my areas of focus in what follows.
Track B1: Divergent metrical structure and alternative lineations

In track A4, I introduced the concepts of alternative lineations and convergent and divergent metrical structure. Lines are the ‘verses in verses’, and alternative lineations involve occasions when the dominant lines are challenged and/or destabilised by alternative possibilities for lineation presented by the rap. Lineation can be implied by many different parameters—linguistic, poetic and musical—and throughout rap’s history, we see the tendency of lines to coincide with the bars of the musical metre (see tracks A4 and A6), resulting in what I have dubbed a convergent metrical structure. In this track, the objects of analysis will be the various common types of divergent metrical structure (those where the lines and bars do not overlap), as well as the ways in which strong alternative lineations tug at the boundaries of convergent (or common divergent) metrical structures to introduce ambiguity or some other aesthetic effect.

I will also present certain essential rhythmic techniques employed by rappers when manipulating lines. I will begin with straightforward examples of a variety of common techniques which depend primarily upon the triangulation of syntax, musical metre and primary rhyme position to manipulate metrical structure. Then, I will look at techniques that blur the boundaries of hypermetric units and rhyme complexes. Lastly, I will present a more extensive analysis of verses by OutKast emcee André 3000 which feature his deliberate application of alternative and ambiguous lineation. In particular, I will focus on some additional linguistic constituents that contribute to lineation as well as the evolution of the dominant lineations as they unfold over time.

**Heteromorphies: Various interactions of primary rhyme, syntax and musical metre**

In his 2008 article ‘Grammetrics and Cognitive Semantics: Metaphorical and Force Dynamic Aspects of Verse-Syntax Counterpoint’, Frank Kjørup explores and outlines a taxonomy of types of *verse-syntax heteromorphies* (or the different ways in which syntax and line segmentation diverge in poetry), likening their interaction to the musical phenomenon of counterpoint. The interaction between grammar and (poetic) metre is the foundation of the branch of metrics Kjørup advocates, known as *grammetrics*. The theory, founded by Donald Wesling, is described via a central metaphor stating that grammar and metre come together like the two blades of a pair of scissors (Wesling, 1996). While Kjørup’s analogy to counterpoint triggers a more phenomenological exploration of the theory he presents, the taxonomy of heteromorphies represents a remarkably sturdy approach to the concept of strong
alternative lineation through the triangulation of specific types of lineation evidence. For Kjørup, these types include (graphic) verse (that is, the line breaks in printed poetry) and syntax, whereas my own analyses drawing upon his taxonomy involve three main axes of virtual reference structures: syntax, musical metre and (primary) rhyme position. These are three of the strongest types of evidence for lineation in rap, and their interaction is the basis of several common techniques which rappers apply to introduce slight divergences in the metrical structure. The first group of heteromorphies described in this track resides fully within the realm of metrical structure—they are examples of divergence or a tendency towards divergence in metrical structure, where either the dominant lineation is misaligned with the metrical span of the bar (a divergent metrical structure) or there are strong alternative lineations which present alternatives to the dominant lineation without actually overpowering or even matching it (a tendency towards divergence). The second group of heteromorphies described in this track, the bridge techniques, involve another structural feature: changes in rhyme class and in the boundaries of rhyme complexes. The reason why different heteromorphy techniques are important to classify and explore is that they are common and can exist within the smallest of hypermetric units (or, rather, symmetrical blocks)—that is, two bars (or a line pair).¹⁰⁰

One motivation for formalising and labelling these techniques is that when discussing passages of rap flows with other rappers, the lack of adequate terminology is often frustrating. There is usually a readily identifiable passage which everyone agrees is really dope, but there are no established terms for talking about the same technique, or similar but slightly different techniques, in another flow. A very interesting aspect of writing this thesis has been that when I ask my rapper friends for examples of specific techniques, they often suggest passages which are examples of slightly different but related techniques, simply because my initial question was not specific enough and there is a distinct lack of nuance in our rap-technical discourse. When I have initiated discussions about these types of techniques in online forums or at informal gatherings, the response has generally been one of interest in and appreciation of the taxonomic description of techniques.

While all heteromorphy techniques represent ways of creating some sort of structural tension at the scissors blades of musical- and poetic-metrical spans, there are clear nuances to their

¹⁰⁰ Note, however, that the bridge techniques (which manipulate the boundaries of rhyme complexes) are especially dependent on or influenced by information from the preceding line pairs or larger symmetrical blocks.
uses and effects. Before embarking on a discussion of their potential overarching cognitive effects or combinational potentials, then, I will describe what they look like and how they are created, then explore some variations upon them.

The first type of heteromorphy, which Kjørup also presents, is well known in the analysis of written poetry—namely, enjambment. Captured neatly via its French etymology, meaning ‘overstriding’, the term typically describes any sort of ‘run-on line’ wherein the syntax continues from one poetic line to the next, ‘striding over’ the boundary of the line ending. Some rap analysts employ this term to describe any type of metre-syntax heteromorphy, in fact. Katz (2008) calls enjambment a ‘mismatch between linguistic and rhythmic constituents’, while Condit-Schultz (2016b, p. 71) writes about ‘effects similar to poetic enjambment’, such as when ‘prosodic/musical phrasing does not align with the syntactic/semantic units’. Like Kjørup, I prefer to be more specific and distinguish between types of commonly used heteromorphies, as they represent very distinct phenomena.

While Kjørup differentiates between ‘run-on lines’ and enjambment, I will not do so here,\(^\text{101}\) as the nuance he captures is tied to line-ending evidence which does not exist in rap flows—that is, the graphic line break itself. Rap’s lack of graphic lineation means that one must specify which lineation evidence assumes this particular role in this heteromorphy. While I do not consider musical metre to be nearly as strong as printed poetry’s graphic line break in terms of lineation (as discussed in track A4), the convention of the line/bar coincidence (that is, the convergent metrical structure) makes it appropriate to consider musical metre to be an analogue to the graphic line break for certain techniques that exist in both rap flows and printed poetry. In rap flows, I use enjambment to denote occasions when (the edge of) the musical metre and the primary rhyme position converge but the syntax diverges and ‘strides over’ to the following bar. In figure B1.1, we see a clear example of this—the primary rhyme falls on beat 4 but the incomplete syntax pushes the flow into the next bar.

\(^{101}\) Kjørup differentiates between run-on lines and enjambment based upon how much force the syntax must use to ‘break through’ the line ending: a run-on line continues effortlessly through the boundary, whereas enjambment describes those occasions when the syntax is temporarily held back by the force of the line ending (Kjørup, 2008, pp. 85-87).
This specific example also exemplifies nicely how lineation in rap verses cannot simply be tied to syntax (in the same manner that it cannot simply be tied to musical metre or primary rhyme position). The syntax of the second line (of figure B1.1, not the verse as a whole) continues for two more full bars (‘I got the cultivating music that be captivating he who listens to the words that I speak as I take me a drink to the middle of the street’), which means that a lineation relying solely on syntax would result in a line spanning twelve beats of the musical background. The primary rhymes, on the other hand, fall predictably and symmetrically on or around the beat 4 positions, establishing a convergent metrical structure, so that the enjambment merely hints at a potential divergence and offers an alternative but non-dominant lineation. This exemplifies, in short, a tendency towards divergent metrical structure.

The next type of heteromorphy is one where syntax and primary rhyme position both diverge from the musical metre, delaying an expected rhyme to beat 1 of the following bar. This technique, called one-rhyme, has many possible variations and is common to most styles of rap flow. In the example in figure B1.2, both the syntactic resolution and the position of the primary rhyme are pushed beyond the boundaries of the musical metre and into the next bar, creating a simple one-rhyme.
The one-rhyme has a particularly strong gravitational pull to it, as it exploits the listener’s anticipation of a rhyme on beat 4 to set up a profound suspension by denying the rhyme’s resolution. Additionally, in this example, it allows the rhyme to coincide with the massively heavy downbeat—‘the One’ with a capital O—invoke by the G-Funk of the musical background. It even introduces a possible alternative lineation, as the secondary rhyme ‘skrittet’/’titte’ (‘crotch’/’peek’) could possibly represent a new primary rhyme. There is only weak support for this possibility, however, because the pre-established hypermetric symmetry sustains the expectation of the primary rhyme, effectively relegating ‘skrittet’/’titte’ to a secondary rhyme. Note also that the line does not necessarily need to end on beat 1 for a one-rhyme to have this strong resolution effect. In the passage in figure B1.2, the syntactical break which coincides with what I consider to be the line ending does not arrive until beat 2 with the word ‘nuts’. This means that, in this particular example, none of the three types of lineation evidence actually coincide. As the music unfolds, we are first presented with the broken expectation of a convergent metrical structure, then a strong rhythmic resolution on the one, then another surprise when beat 1 turns out not to be the line ending.

While a simple one-rhyme like that in figure B1.2 introduces an unexpectedly large gap between rhyme instances, another technique which introduces slightly less built-up tension is the half-line rhyme to one-rhyme, as exemplified in figure B1.3. Here, the lyrical line is naturally subdivided into two parts via the addition of one extra primary rhyme instance at what would be the halfway point of the line if the metrical structure were convergent. Whereas the decrease in rhyme density and the huge build-up of tension from the delayed resolution in a simple one-rhyme comprise a very strong metrical statement, the addition of the half-line rhyme (and its accompanying weak alternative lineation) is somewhat less so. The delayed resolution of the one-rhyme introduces less tension when it is more proximate to the previous iteration of the rhyme instance. As listeners, our itch for the rhyme has already been scratched (even if the rhyme arrives early), so when the one-rhyme arrives, it represents a subtler type of shift from convergent to divergent metrical structure.
The subdivision of a line into two equal parts to introduce a competing lineation represents a very common technique in poetry which Fabb (2002, p. 149) characterises as ‘mid-line boundaries that resemble end-of-line boundaries’. In the case of the half-line (or ‘mid-line’) to one-rhyme technique, this competing lineation is strengthened (or at least retained as an alternative) when the primary rhyme and end-of-line boundary are forced beyond their expected position after beat 4. Whereas one might begin to wonder where the line was headed while awaiting the resolution to a simple one-rhyme, the half-line to one-rhyme instead begs the question of whether the line ended before it was supposed to. Note too that there are ways of subdividing lyrical lines other than via an increase in rhyme density, and as an alternative weak lineation, the extended reinforcement of the half-line point is an effective technique.102

The dual-planed musical metrical frameworks of many trap-derived or trap-inspired hip-hop tracks (see track A6) has some interesting implications for one-rhyming and the half-line alternative lineation, as exemplified in figure B1.4.

---

102 A switch to the halfway point of the musical metre as the dominant lineation in an extended section of a flow is a type of divergent metrical structure which blends well with surrounding passages of convergent metrical structure, as can be heard and seen in the example of ‘Anaconda’ in track A4 (page 68). It is similar to a switch between on- and off-beat phrasing, in that it is the expected pattern but shifted one position over in a subdivisional layer.
This example showcases a type of trap subgenre wherein the two alternative tactus tempi are virtually equal. ‘Latin trap’ superimposes a salsa groove onto the slow boom-bap typical of trap, and the two tempi wrestle for dominance throughout the track; a flow can be structured to emphasise one or the other and/or switch between the two. This is what happens during Bad Bunny’s verse (starting at 1:24). Initially, his lines are long, coinciding with the bars of the slow trap timekeeper. Then, he switches his density referent from the sixteenth-note layer to its triplet version and begins emphasising the half-line boundary/salsa timekeeper by using pauses and increasing his rhyme density. At this point, the listener is likely to experience the quicker tempo (which is closer to ‘maximal pulse salience’ discussed in track A6) as the tactus, converting the passage in figure B1.4 into a simple one-rhyme because there is no rhyme instance in the entire bar, and the anticipated rhyme is delayed to the following beat 1. However, there is a complication. While the most common position of the primary rhyme in the rap corpus is on beat 4, this is no longer the case in the ‘double-time’ tempo here. As the figure shows, the first instance of the primary rhyme (‘Plaza’) lands on beat 3 rather than beat 4 of the salsa timekeeper. This means that the typical rhyme position in a trap metre is consistent with the slow tempo and the extensive half-line division, not with the fast tempo. Still, this does not have to mean that the slow tempo is the definite tactus for the listener. Instead, it simply implies that there are alternative typical rhyme positions depending upon which tempo one is entrained to. ‘Plaza’ is not ‘early’ because it lands on beat 3 rather than beat 4 of the salsa timekeeper; on the contrary, it is completely genre typical. In the end, the one-rhyme in figure B1.4 falls somewhere between a simple one-rhyme and a half-line to a one-rhyme.

A one-rhyme does not have to be a delay of an expected rhyme—in the case of an added one-rhyme, a rhyme still lands on the expected position but is immediately followed by another rhyme instance of the same rhyme class. Depending on its interaction with other lineation evidence such as syntax, this can elongate a line after it was thought to have been ended, or it can represent the start of a new line (see figure B1.5).

---

103 This is also the case regarding the double time version of ‘a lot’ as presented in track A4 (page 65) and reiterated in the slow boom-bap tempo in track A6 (page 104).
As a heteromorphy, the added one-rhyme represents an interesting example of a dominant lineation which is established and immediately challenged, requiring, in turn, fresh reassessment. In the example in figure B1.5, the lineation evidence coincides and presents an obvious line-ending after beat 4, implying the expected convergent metrical structure. However, there is no pause in the flow, and the primary rhyme (which can be considered a split rhyme in this context)104 is immediately reiterated. The first intuitive interpretation is that ‘you tend to be gross’ is the start of a new line, but because the syntax is closed and Diaz inserts a pause in the flow, the new line ends as soon as it begins, spanning only a pickup and a single beat. This line is too short to be viable as its own segment, so the listener wants to attach it to something; the only realistic option is the preceding line, to which it is connected by rhyme.105 There is some structural impetus to allow the short line to remain unattached, in that the preceding line is subdivided by half-line phrasing. In that line, however, the rhyme placement—with the first rhyming syllable on the beat and the second one syncopated—forces the subdivided half-lines to span two beats, which makes them categorically different from the short line, even though they are similar in absolute time.

There are many variations of the one-rhyme technique which are fairly similar in the way they sound. For instance, if syntax and musical metre converge but the primary rhyme position is

---

104 The final stressed syllable of the multisyllable rhyme rhymes with the following final punchline rhyme (see fn.105), and both stressed syllables (the first and the last) rhyme with the preceding line (‘bending your hoes’, ‘spending on clothes’). The unstressed middle syllables do not rhyme consistently with the previous line or within the line in the figure, however.

105 The verse is an example of extended monorhyme, but there are variations like the split rhyme from figure B1.5 and its preceding line. The (too) short line of ‘you tend to be gross’ has a rhyme connection to the line that begins right after it, which ends with ‘chose’, so it could conceivably be interpreted as part of that line (which would introduce a convergent metrical structure). Because the first part of the split rhyme is lacking that rhyme connection, however, it is much weaker than the connection it has to the preceding line.
delayed from the expected beat 4 to the following beat 1, this is still a one-rhyme, but the rhyme might not imply a dominant lineation; the triangulation of syntax and musical metre might well have stronger lineating power than the primary rhyme position on its own (or in conjunction with weaker types of lineation evidence such as other parallelisms or breathing pauses). In the case of added one-rhymes, the syntax can overlap with either musical metre or rhyme position, or with neither, or even with both. We see examples of all these variations in figure B1.6.

![Figure B1.6: One-rhyming in the first verse of Ms. Lauryn Hill’s ‘Doo Wop (That Thing)’ (1998, 0:35). Metrical structure of lines/bars 4–10, primary rhymes marked in red; added one-rhymes in lines/bars 13–14 and line 21 (bars 21–22), syntactic boundaries in blue brackets.](image)

Beginning with lines 4–5, there is a simple added one-rhyme where the syntax is slightly ambiguous. ‘To begin’ can be considered part of either line, offering two alternative lineations through its syntax alone. Because there is a primary rhyme instance on beat 4, however, the dominant lineation is the convergent one. The added one-rhyme and the syntactical ambiguity therefore present a fairly strong alternative lineation, creating a tendency towards divergence within a convergent structure.

In lines 7–9, there is repeated one-rhyming which creates a consistently displaced metrical structure, a symmetrical type of divergent metrical structure (as in the example of figure A4.7, page 80, which features convergent poetry but a divergent metrical structure). Here, there are
no rhyme instances on beat 4, so both rhyme and syntax indicate a dominant lineation which is not aligned with the musical metre.

Lines 13–14 display clear ambiguity along the syntactical axis, as ‘girlfriend’ can be associated with either line or both at once (even though it is not actually syntactically part of either, strictly speaking), thereby suggesting an alternative lineation in the form of an *elision*. Note that I use ‘elision’ in its musical sense here, meaning that the final musical unit of one phrase is simultaneously the first unit of a new phrase.\(^\text{106}\) In linguistics, the term has a different meaning (referring to sounds being deleted from words in order to shorten them). Like lines 4–5 (and 21–22), the one-rhyme is the added kind, so there is a rhyme instance which implies a convergent metrical structure.

As an interesting side point, Ms. Lauryn Hill does not seem particularly invested in emphasising symmetrical blocks of four or more bars in this verse. Its metrical structure varies between convergent and (slightly) divergent, but none of the long stretches of consistent variation (such as bars 7–10) aligns with the common structural divisions of a rap verse.

The aesthetic effects of one-rhymes derive from the tension and release, and from the ambiguity among alternative dominant interpretations of a flow’s structure. These parameters, however, rely on a lot of nuance. To reiterate a central point, for example: a rhyme is not yet experienced as a rhyme until the second rhyme instance appears, meaning that the first instance of a rhyming pair or chain always receives its rhyme status *retroactively*. For techniques such as one-rhymes, this has implications for which types of expectation inform the tension of a delayed rhyme. If the primary rhyme is already established (it is part of a longer chain/a larger rhyme complex), there is a correspondingly stronger expectation of a resolution: one expects a specific rhyme. If the one-rhyme is the rhyme instance which establishes a rhyme connection, one should theoretically *not* be expecting it. However, the convention of the line/bar coincidence makes us expect some kind of rhyme on beat 4, so there is definitely *some* expectation in any case—it is simply different, both less explicit and more vulnerable to being subsumed by secondary rhymes or other alternative lineation evidence. The degree of expectation and tension is similarly impacted by where the one-

\(^\text{106}\) Very much in the spirit of this thesis, the German word for (musical) elision is *Umdeutung*, meaning ‘reappraisal’, because it entails a musical unit being first experienced as one thing, then reappraised as another (or, really, as both). This notion is mostly used in the analysis of harmony, but I find it evocative in the analysis of rap flows as well.
rhyme is placed within the framework of a verse. Consider the different variations in figure B1.7 of simple one-rhymes with no complicating factors.

**Figure B1.7:** Different hypothetical placements of simple one-rhymes within groups of bars (or ‘hypermetre’). Symmetrical hypermetric blocks marked with green (line pairs) and blue (blocks of four bars).

In example 1, the one-rhyme is unprepared (that is, the one-rhyme establishes the rhyme), but since we expect a conventional line-pair with a convergent metrical structure, we eagerly await the resolution of the one-rhyme. In example 2, however, the first line pair is resolved, and we do not necessarily expect the rhyme complex to extend any further. Thus, the one-rhyme is not anticipated to the same degree as it was in example 1, even if the rhyme has already been established. Perhaps the strongest expectation is created in example 3 because the extended rhyme complex is established in the third bar, and we expect a specific rhyme class to round off the symmetrical four bar block. The opposite is in effect in example 4, where a fully symmetrical rhyme complex converges with a block of four bars, so there is little expectation of the rhyme complex carrying on. In short, one-rhymes interact with the *gravity of the form*. Different means of emphasising larger structural blocks and the full verse form will be the topic of the next track; in what follows, we will look at how rhyme-class changes and the boundaries of rhyme complexes interact with and influence the metrical structure.

---

107 Note that I only use ‘hypermetre’ for symmetrical groups of bars. The term does not imply that ‘hypermetric units’ are experienced in a metrical way.
Bridge techniques

In the examples throughout this track, we have yet to encounter techniques wherein the rhyme class changes. It is, of course, common that a single primary rhyme ‘rules’ an entire hypermetric unit with rhyme complexes varying in length from a simple line pair to a block of four bars or (using an extended monorhyme) even an entire verse. One way in which rappers destabilise this structural predictability is by employing bridge rhymes to create connections within or across hypermetric units, introducing structural ambiguity in the relationship between metrical structure and rhyme complexes.

Alim (2003) uses the term ‘bridge rhyme’ and ‘bridge rhyming techniques’ to describe the ‘next-level poetry’ of Pharoahe Monch and his crew:

the bridge rhyming technique (…) is widely used in Hip Hop poetics. In musical terms, a bridge is a transitional passage that connects two subjects or movements. The important point here is that the bridge is transitional, allowing two seemingly disparate or distant objects/movements to coalesce into one. Pharoahe and his crew of Hip Hop collaborators use this bridge rhyming technique in a variety of ways to form a continuous highway of rhymes that connects two seemingly ‘unrhyymable’ words/verses. The smooth transition, when done skillfully, is noticeable only on a subconscious level. (Alim, 2003, p. 75)

In the analysis of the metrical structure of rap flows, the notion of ‘coalescing’ and the image of ‘a continuous highway of rhymes’ suggest other methods through which rappers might play with the boundaries of the different constituent blocks within a rap verse. As we will see (and the attentive reader will already have seen throughout the previous tracks), the boundaries of rhyme complexes need not follow the boundaries of any other natural segmentation of a verse, such as symmetrical groupings of bars or dominant lineation. In fact, rhyme complexes are not particularly territorial, and there are often widespread border crossings and cross-pollinations between them. To complicate this situation even further, the less structurally significant secondary rhymes might also be used as bridge rhymes, connecting sections in ways which are smooth and skilful enough that they, as Alim writes, will be ‘noticeable only on a subconscious level’.

Perhaps the most common form of bridge rhyme involves the situation wherein the first line of a new section (most often a new line pair that may or may not be part of a new larger symmetrical block) begins with a rhyme from the preceding primary rhyme complex but ends
with what will become the primary rhyme of a new rhyme complex. One variant of this involves the use of a simple added one-rhyme, as in figure B1.8.

**Figure B1.8:** Simple added one-rhyme bridge in Big Boi’s verse on OutKast’s ‘Ms. Jackson’ (2000, from 1:03). The two primary rhyme classes are marked with red and blue outlines, respectively. The same colours show the rhyme complexes within the convergent metrical structure in the bar representation to the right.

While this passage retains its convergent metrical structure, the bridge rhyme creates a heteromorphy between the metrical structure and the rhyme complexes, which helps to soften the edges between the hypermetric units. A very similar example is presented in figure B1.9, where, in place of a one-rhyme, a bridge rhyme occurs at the half-line point. This even subdivision of the line makes the weak alternative lineation presented by the rhyme position and syntactical subdivision (seen with a comma in figure B1.9) slightly stronger than it was in the previous example.

---

108 Hirjee and Brown (2010, p. 123) call these types of bridges (whether one-rhyme bridges or half-line bridges) ‘link rhymes’ and reserve ‘bridge rhyme’ for rhymes where both rhyme instances are not ‘line final’ but instead ‘internal’. They do operate within a framework whereby the ‘line’ is derived from a convergent metrical structure, however.

109 There is ample reason to consider any added space between rhyme instances to be a strengthening factor for the rhyme. Both Ohriner (2016a, pp. 158–159) and Komaniecki (2019, pp. 71–72) note that factors such as a ‘lower limit of proximity’ to the perception of a rhyme might influence our subjective interpretations of rhyme. If two rhyming syllables are adjacent to one another, they will most likely not be perceived as rhyming. Metrical (and other structural) contexts influence our experience of rhyme just as rhyme influences our experience of structure.
Figure B1.9: Bridge rhyme in the first verse of South Park Mexican’s ‘High so High’ (1999, 0:36–0:47). Primary rhymes indicated by coloured boxes, line pairs with brackets. Bridge rhyme emphasised in bold red letters. Rough beat positions marked in green.

Note that while the bridge rhyme is at the half-line point, it falls on beat 1 of the musical metre. This is because this passage has a consistently displaced metrical structure like the repeated one-rhyming in ‘Doo Wop (That Thing)’ (figure B1.6). In this case, however, the lines begin with a one-beat pickup, making them stretch from beat 4 in one bar to beat 3 in the next. Thus, the ‘half-line rhymes’ fall on beat 1, while the ‘end rhymes’ fall on beat 3. Note that I do not consider this type of structure a one-rhyme variant even when the rhyme is placed on beat 1. One-rhyme techniques are characterised by the rhyme being postponed beyond or added after the expected rhyme on beat 4, and that is not the case in this example.\footnote{One can also argue, then, that extended repeated one-rhyming (like Ms. Lauryn Hill’s in figure B1.6) stops having a one-rhyme effect after a couple of repetitions. The passage in figure B1.9 is below that threshold, however, and the consistent one-rhyme displacement does not quite manage to establish itself as a ‘new normal’.}

Another type of bridge rhyme appears in figure B1.10, where we see an added one-rhyme which is not fully a rhyme but instead partly rhymes with a preceding primary rhyme and partly with the following one.\footnote{This is related to but fundamentally different from the ‘rhyme shift’ (Komaniecki, 2019, pp. 70-71) discussed in the rhyme dictionary in track A5. The gradual ‘twisting’ of vowel sounds in a rhyme shift is more subtle and results, as Komaniecki points out, in all the rhymes being considered part of one single rhyme complex. A pivot rhyme is clearly heard as something different from at least one of the implicated rhyme complexes but also as connected to both of them nevertheless.} I call this technique a pivot rhyme.
Figure B1.10: Pivot rhyme in the opening of Side Brok’s ‘Setra’ (2004, 0:12–0:22). Primary rhymes marked with red and blue outlines, respectively. Pivot rhyme in bold red and blue letters. Musical beats and bar structure marked with green numbers and dotted line. Asymmetrical syntactic unit (and dominant lineation) marked in yellow bracket and dotted line. Consonance and assonance connections between primary rhymes and the pivot rhyme indicated at the bottom of the figure.

The rhyme connections between the pivot rhyme and the primary rhymes are necessarily weaker than the internal connections between the primary rhymes within their respective rhyme complexes (otherwise, they would rhyme with one another and form a single rhyme complex instead). Still, the connective effect is clear (particularly between the first primary rhyme and the pivot, with 100 percent consonance/alliteration), and we get a sort of rhyme complex elision, with the pivot rhyme being part of both rhyme complexes and neither of them at the same time. The enjambment (in the syntax but also emphasised by the pivot rhyme and the pause following it) is also an indicator of the structural importance of the word containing the pivot rhyme, and of the resulting dominant lineation. Here, the emerging alternative lineation which develops as the rap unfolds is crucial to the aesthetic expression, where the convergent expected line ending of ‘Kreta’ is immediately contradicted by the following words. We will come back to this type of gradual reinterpretation of dominant lineation shortly.

An exceedingly rare variation involves the combination of an early onset primary rhyme instance with a tacked-on bridge rhyme at the four, as exemplified in figure B1.11. One of the reasons this is a quite harsh bridge rhyme technique is that, in addition to breaking many

\[\text{112 As the case would be with a ‘rhyme shift’. There is another speech-sound parallelism between the two instances of ‘ret’ in ‘Kreta’ and ‘Retsina’, but it is not a rhyme.}\]
structural expectations, the structurally redeeming new primary rhyme is, like all rhymes, first heard only when the second rhyme instance arrives. Thus, the prominent placement of the first instance of the bridge rhyme is not ‘justified’ until the end of the following line/bar.

The bars transcribed in figure B1.11 provide a good example of the structural weight of syntax and musical metre (and the expected primary rhyme on beat 4) because even though the lines are littered with secondary rhymes, the surprise of the line ending with ‘ego’ is striking. Similarly, one must undertake a major reappraisal of the flow’s structural intricacy when the status of ‘ego’ as the new primary rhyme is finally revealed. While this type of variation, with the asymmetry of three bars housing two rhyme complexes, is quite uncommon, the structural parameters which are manipulated are the same as ever: rhyme position, rhyme change and syntax.

Figure B1.11: Early primary rhyme onset and bridge rhyme in the first verse of OutKast’s ‘Skew It on the Bar-B’ (1998, 0:20–0:26). Primary rhymes indicated by coloured boxes, secondary rhymes underlined. Note that the light blue secondary rhyme is considered to be separate from the blue primary rhyme because the primary rhymes connect to the following line/bar, whereas the rhyme ‘RICO’ and ‘Chico’ does not have the added /ɔ/-syllable.

This passage also hosts a fascinating example of how a listener’s knowledge of pop-cultural references in the lyrics can affect the experienced lineation. Unless one is familiar with the gaming console Colecovision, the dominant lineation is clearly convergent. However, if one is a gaming console-afficionado, the syntactical boundary (and the resulting enjambment) might well influence the interpretation of dominant lineation, resulting in a divergent metrical structure with the line ending after ‘Colecovision’ in the fourth bar.
While we have discussed at length rhyme’s ‘lopsidedness’, in that the first instance of a rhyming pair does not actually rhyme until its partner rhyme instance reveals it, I find it fascinating how expected convergence muddies these waters at times, as in ‘Skew It on the Bar-B’. The first instance of the rhyme is placed where the primary rhyme is expected to be, so we are more likely to expect ‘ego’ to rhyme with something than any other stressed word in these bars. My favourite example of this type of delayed resolution of a rhyme is from my bandmate Bennærn. It is in the opening of the first-ever Sinsenfist track, and the first-ever Sinsenfist verse, and it blew my mind when I first heard it. The opening bar goes ‘Når jeg detonere, da rykker det ettertrykkelig’, and when the line continues in the following bars, there is no rhyme to this line in sight: ‘. . . i rockefoten, men ikke si det kom som no’ sjøkk / for jeg er hot som fäkk, men gjør deg fæn ikke . . .’. Throughout these bars, the /ɔk/ (and /ɔt/)-rhyme has taken over, so you have almost forgotten that the opening statement of bar 1 ends with ‘ettertrykkelig’ when suddenly, at the end of bar 3, it is revealed as a rhyme with the word ‘lykkelig’ landing on beat 4. This primary rhyme is even reinforced and reiterated in the next bar, ending on ‘sykler’ (Sinsenfist—‘Spreng’, 2014). This blurring of lines, rhyme complexes and what constitutes a primary and a secondary rhyme as the first four bars unfold is to this day one of my favourite pieces of Bennærn flow.

To round off the presentation of bridge rhyme techniques, I will present an example (figure B1.12) of a combination of different bridge rhyme variations used within a single flow.
Figure B1.12: Excerpt from the first verse of Nas’s ‘N.Y. State of Mind’ (1994, 1:29–1:52) showing metrical structure, bridge and pivot rhymes as well as partial overlaps in the primary rhyme (dark blue to light). Black dots under transcribed lyrics show beat positions, crosses show rhyme positions within metrical structure, coloured lines show secondary rhyme instances.

This verse excerpt demonstrates that the interaction between the boundaries of rhyme complexes and metrical structure can be just as intricate as that between bars and lines. Nas’s extensive use of secondary rhymes and partial rhyme connections to and between primary rhymes actively blurs the edges of his rhyme complexes, making the flow appear to simply cascade forward.

In the first two lines, the divergent metrical structure is supported by both primary and secondary rhymes. The initial primary rhyme is presented as three rhyme instances on three
successive beats with the line ending (according to our expectations) prematurely on beat 3. When the next line introduces the /-iggas/ rhyme, it appears that a new rhyme complex has been established, only for it to be relegated to secondary rhyme status by the return of the /-em/ primary rhyme. Not only is the previous rhyme complex reinstated but the rhythmic and poetic gesture of three repeated rhyme instances on three successive beats is reiterated, creating another layer of parallelism and pushing the final rhyme instance into the third bar as a one-rhyme. The lineation is highly ambiguous, however, as the syntax allows for a line ending after ‘name’ on beat 4. In addition, there is no line ending after the final rhyme instance of the rhyme complex, and the clause ‘and claim some corners’ could very well be interpreted as a short line of its own. That interpretation is validated by the following short line ‘crews without guns are goners’, which establishes both the new /:n̩əs/ rhyme complex and the beginning of a section of convergent metrical structure. The multiple alternatives presented for line endings from the end of the second bar to the middle of the third bar prompt a chain of reappraisals, complicating the establishment of a clear dominant lineation. The lineation I have chosen to use in figure B1.12 must be considered strongly subdivided—an average, more or less, of the emergent and reappraised lineation a listener might experience. The third bar, for example, could be reappraised as one subdivided line in a convergent line pair, but for it to be the dominant lineation, we would have to ignore the strong evidence of the repeated three-primary-rhyme-instance figure. If anything, both lineations might be experienced as dominant (but overlapping).

The mosaic three-syllable rhyme class dominating the rhyme complex of lines/bars 5–6 in figure B1.12 spans two beats, making it possible for the rhyme instance to fall both on beat 4 of bar 6 and be a one-rhyme because the second stressed syllable anticipates beat 1 of bar 7. The rhyme techniques in the rest of bar 7 and the connection to bar 8 are the most spectacular gestures here, however. The new primary rhyme (/ak/) is introduced as a pivot rhyme, with ‘block’ having perfect consonance (that is, the same consonant sounds) with ‘black’, the final part of the previous primary rhyme. However, there is a secondary layer to the pivot rhyming, as the word ‘snitch’ in bar 7 has a similar perfect consonance connection to ‘snatch-’ in bar 6. This double pivot of ‘snitch on the block’ sets up a kind of mosaic rhyme connection to ‘niggas knocked’ as well, even if the latter is not truly of the same rhyme class. The two rhyming syllables between the two groups are separated by a different number of unstressed syllables and are part of different rhythmic figures (that is, they do not display a perfect
rhythmic rhyme). Thus, I will analyse them as two secondary rhyme instances preceding two primary rhyme instances, even if they do set up a more holistic rhyme connection than that indicated by these terms. The primary rhyme class of /ak/ retains its role as the final syllable of the primary rhyme throughout the rest of the passage in figure B1.12, though the one-syllable rhyme class becomes a three-syllable one in the final stage of the gradual rhyme evolution of the section’s primary rhymes.

The long lines of André 3000

The interplay between bars, lines, rhyme complexes and various types of evidence for lineation is evidently central to the more or less advanced flow techniques employed by rappers of many different stylistic, cultural, historic or coastal backgrounds. The final section of this track will be devoted to a representative of the ‘third coast’ of American hip-hop, an artist renowned for his idiosyncratic flow style: OutKast’s André 3000. We have already gotten a slight tease of the two OutKast emcees’ propensity for divergent metrical structure and ambiguous lineation in the first part of this track, and it will be clear that André 3000 is a master of playing with the listener’s expectations of syntactic closure, convergence between metrical spans and other emergent parallelisms in many of his verses. The analyses will involve parts of different tracks from the group’s third album, Aquemini, from 1998 and focus on how his flow keeps suggesting line endings that in immediate hindsight are not the dominant ones. To achieve this sleight of hand, André 3000 takes full advantage of the triangulation of syntax, primary rhyme position and musical metre, as well as well-timed pauses in his flow which either nudge the listener towards a specific dominant lineation or create additional ambiguity.

Figure B1.13 presents a segment of André 3000’s verse on ‘Rosa Parks’ (the song’s second verse, after Big Boi raps the first), where he uses the placement of primary rhymes at the predictable position of beat 4 to evoke a sense of line breaks even as the syntax and semantic content (as well as a carefully placed secondary rhyme) make us repeatedly reconsider what the dominant lineation is.

---

114 One interesting and important nuance to this is that the rhymes are between two pairs of successive stressed syllables, even if the rhythmic surroundings are otherwise dissimilar.

115 American hip-hop’s ‘third coast’ is the South. While hip-hop culture in the nineties had a clear East Coast/West Coast divide, the development of a new style of hip-hop in Southern cities like Houston, Memphis and most notably Atlanta was hugely important to the stylistic trends of hip-hop in the twenty-first century.
The rhymes are placed consistently and precisely on the metrically strong positions, with the initial /əʊ*/ sound landing on beats 1 and 2 of bar 15. The rhyme evolves into a two-syllable /æ*/-/*æ*/ (have to float / that boat / last quote) which is still metrically strong on the eighth-note positions,\(^{116}\) and these insistent eighths strongly argue for a line break producing (an expected) convergent metrical structure. The sentence ‘and I hope I never have to float in that boat’ readily comes across as a complete statement, and an initial logical interpretation is that the line ends after ‘boat’. However, as the flow continues, the next three words ‘up shit’s creek’ is clearly (syntactically and semantically) part of the preceding line. The rhyme and musical metre which implied a line ending were mere deception, and we are forced to reinterpret what the dominant lineation is.

The next two words—‘it’s weak’—are ambiguous in terms of their status. The syntactical unit of the previous line is closed after ‘up shit’s creek’, but there is a clear rhyming connection between ‘shit’s creek’ and ‘it’s weak’ which pulls them together. Since the two syllables of ‘it’s weak’ are too short to represent a line of their own,\(^ {117}\) our first inclination might be to tack them on to the preceding line—connected by rhyme, that is, if not syntax. André 3000 also does not pause between ‘creek’ and ‘it’s’, so there is no indication of a line break in the

\(^{116}\) As discussed in track A6, eighth notes are relatively metrically strong in a context in which the density referent is at the sixteenth-note level or higher.

\(^{117}\) Fabb includes ‘size’ as one of the many types of evidence for lineation: ‘While there can be short lines and long lines, there are normative expectations relative to the tradition that lines will be a certain length, which means that evidence for a line boundary will be strengthened if it falls close to the normative length of the line’ (Fabb, 2002, p. 144). A very short two-syllable line is not congruent to the surrounding lineation structure in the flow analysed here, so this potential interpretation of lineation is significantly weakened.
performed delivery but rather the opposite. Again, though, the unfolding flow indicates a clearer lineation. ‘It’s weak’ is part of the next syntactical unit: “it’s weak” is the last quote’, a line which again seems to end convergently, with the primary rhyme falling predictably on beat 4 and a syntactic unit which could be interpreted as closed. In the exact same manner as the previous bar, we initially expect the line to be ending before it is revealed that it continues into the next bar with another long line: “it’s weak” is the last quote that I wanna hear when I’m going down’.

In these lines, André 3000 shows how a performed emphasis on the rhythmic figures and positions of rhymes, combined with the expectation of a rap verse’s typically convergent metrical structure and symmetrical grouping of bars, can present strong alternatives to lineation that, as the flow emerges, keep us busy reinterpreting the poetic and musical structure. Even though primary rhyme position and musical metre coincide in these lines, the syntactical and semantic information in the lyrics is strong enough to imply a dominant lineation\(^\text{118}\) in the face of this conflicting evidence. What ‘dominant lineation’ actually entails in a passage such as this is also an interesting question because the ambiguity—and the constant reinterpreting of where the line ends—is an essential part of the flow as we experience it, and André 3000 manipulates our expectations masterfully. Whatever dominant lineation is present is fundamentally defined through its relationship to the alternative lineations it defeats.

While the expressive timing in ‘Rosa Parks’ never strays far from relative isochrony (the durations of most beat- and subdivision-level rhythmic events align closely to the canonical time-value classes) and the consequent clarity in (quantised) musical rhythmic structure, the verse we turn to next—André 3000’s first on the title track of Aquemini—is characterised by a much looser adherence to isochronous subdivision. His expressive timing evokes a certain speech-likeness, even if the phrasing, pauses and variations in tempo are not evocative of actual speech but rather comprise a stylised imitation of some aspects of spoken language. The timing obfuscates the specificity of the isochronous beats and musical metre of the musical background—a background which is already quite sparse in its categorical rhythmic information. There is no established duration(s) or swing ratio for sixteenth-note subdivision

\(^{118}\) This interpretation resonates with Adams (2020), who identifies syntax (and rhythmic motive) as a phrase-defining parameter in hip-hop. There is some overlap with Adams’s conception of a musical phrase in hip-hop and poetic lines, but ultimately, they are quite different things, as Adams delimit ‘phrase’ to passages that ‘contradict our normative expectations’, while ‘lines’ are all the poetic spans in a piece of poetry (like a rap flow).
save for what is present in the vocals and the cuts (DJ scratching) in the beginning, and André 3000 takes advantage of this unmarked subdivision canvas to perform a rhythmically expressive flow. Both his delivery, his timing and (as we will see) the lyrical content work together to give the impression of free association or automatic writing in the verse. But the parameter that really hammers this feeling home is André 3000’s use of divergent metrical structure and ambiguous lineation, as can be seen in figure B1.14.


Interestingly, this is not the case in André 3000’s second verse on the same track—one which many considered to be one of his finest, and which cemented his reputation. That verse is characterised by a very high rhyme density using a repeated multisyllable rhyme rather than the ambiguous expressivity of the verse analysed here.
Divergent metrical structure is dependent on convergent metrical structure’s existence as a reference, and even in verses like this one, with the extensive use of divergent metrical structure, it is rare that a convergent metrical structure is not at least hinted at. When establishing this reference in the opening bars of his verse, however, André 3000 leaves it at a bare minimum. The first two lines are not connected by rhyme at all; there is only an internal rhyme connection within each line, symmetrically subdividing them into two, rhyming on beats 2 and 4. Then, having adhered to (and commented on) a traditional convergent metrical structure, André 3000 departs on a winding, exploratory path in the next part of the flow.

The third line, as it does not connect through rhyme with the previous lines, presents its first indication of a potential line ending with the sentence concluding ‘dirty dollars’. However, instead of starting a new line, André 3000 start reeling off other constituent parts of his ‘home of the brave’, first both rhyming and repeating the same four-note (and four-syllable) rhythmic figure (‘beauty parlors’ and ‘baby ballers’), then hinting at a departure by breaking the rhythmic repetition (‘bowling ball impalas’) and finally trailing off by closing the syntactical unit without either rhyming or adhering to the boundary of the musical metre (‘street scholars majoring in culinary arts’). The rhymes are not placed in any consistent metrical position and are delivered via a very loose timing between the beats of the musical background and the musical metre (the crosses in figure B1.14 are approximate positions within the line rather than strict beat positions). At this point, then, the listener has been led off the beaten path, and André 3000 has established that he does not care to be shackled by conventional rhyme placement, consistent line length or defined musical metrical spans.

Continuing with what I have marked as lines 4 and 5 in the figure, André 3000 keeps presenting alternative lineations, blurring the very concept of a dominant lineation altogether. The /əʊ/ rhyme is only weakly presented as a primary rhyme via the exaggerated pause after the word ‘dough’, as the first instance of the rhyme—‘know’—is not placed where we would expect a rhyme to be. In other words, we do not encounter ‘know’ as a potential rhyme instance until that rhyme is revealed by ‘dough’ (and the subsequent pause), and this one-syllable rhyme, placed away from beat 4, is weak enough that the secondary rhyme of ‘scratch’ and ‘catch’ makes us question whether the /əʊ/ rhyme is even particularly

---

120 The connection between ‘baby ballers’ and ‘bowling ball impalas’ (and the preceding ‘beauty parlors’) is also strengthened by alliteration, which helps alleviate the weakening of the rhythmic parallelism. Another alliteration shows up in ‘street scholars’ as well.

121 Note that the first two syllables of ‘culinary’ can be thought of as part of the rhyme complex. The main verbal stress on the third syllable disrupts the feeling of rhyme, however, weakening it significantly.
significant. Only when it is revealed to be part of a long rhyme complex with ‘bought’, ‘talk’ and ‘walk’ (the latter two rhyming with the former through André 3000’s thick Atlanta accent) is the real significance of the line-initial ‘know’ established once and for all. One can, and perhaps should, argue that dividing these lines through syntax, as I have done in figure B1.14, is not necessarily the ‘correct’ dominant lineation as the flow unfolds. As in ‘Rosa Parks’, the introduction of the secondary rhyme creates a sense of connection between ‘scratch’ and ‘but see the catch . . .’, and the pause after ‘dough’ is also an argument for moving the phrase ‘from scratch’ to the following line. My reason for the lineation is that there is a clear pause between ‘scratch’ and ‘but’, and there is a phonological hint of finality in the way André 3000 raps the word ‘scratch’. Ultimately, that is, the ambiguous lineation has a clear aesthetic effect, whichever way one interprets it.

Continuing to lines 7 and 8, the dominant lineation is now clearer but still constantly challenged by various weak alternatives contradicting it. First, the bridge rhyme ‘walk’, bridging the /əʊ/ rhyme complex via a simple added one-rhyme in the line dominated by a new primary rhyme (‘halfway’ rhyming with ‘pathway’). This bridge rhyme is emphasised by André 3000 pausing slightly in the middle of the syntactical unit, as though he wants to make sure the listener catches the rhyme. ‘Halfway’, not yet a rhyme when it appears (it needs a rhyming partner) is still experienced as line final through the syntactic closure and André 3000’s slight pause. It also falls on the heavy downbeat of beat 1, establishing what turns out to be a repeated one-rhyme line-pair when ‘pathway’ is placed on the next beat 1, exemplifying why a repeated one-rhyme structure is what Tsur would consider to be convergent poetry (see track A4, page 80). This structure, however much it is divergent metrical structure, has a ‘clear-cut shape’, is symmetrical, and emphasises the musical metre, if not by rhyming on the fourth beats then by rhyming on consecutive first beats.

In the two next lines, André 3000 pushes the rhyme one step further, to the anticipated syncopated beat 2, before starting an interesting trail towards concluding the verse with convergent metrical structure. The segment of the flow in figure B1.15 contains several short utterances which are both difficult to classify as lines themselves and to confidently combine into longer lines.

In particular, the four-word phrase ‘more than a hobby’ is ambiguous in terms of whether it belongs to either the preceding or following line or should be considered a (very short) line in itself. The rhyme connection goes both to the preceding line, beat and bar (‘probably’), and to the following line, with the word ‘wobbly’ on beat 3. As the potential lineations arise in turn, one might initially consider ‘more than a hobby’ to be a prolongation of the preceding line, then connect it to the following line when ‘wobbly’ appears, then reinterpret it again when the line introduces a new primary rhyme on the strong beat 4 (indicating that the foundation of the rhyme complex to which ‘hobby’ belongs is situated in the previous lines rather than the current one).

The high rhyme density and rapid change of rhyme class continues, with ‘all’ and ‘y’all’ being succeeded by ‘drift’ and ‘gift’ within the next bar. The connections from short utterance to short utterance keep contradicting one another, as the unit which makes the most sense to connect semantically—‘I’m sorry y’all, I often drift’—is split down the middle by a change in rhyme class. The next short statement—‘I’m talking gift’—is performed in one phrase with the following line, introducing the expectation that there will be a rhyme connection between ‘gift’ and the line ending. That expectation goes unfulfilled, however, with the introduction of yet another rhyme class when ‘grill’ ends the line (marked by the syntactic break, a pause in the flow and an adherence to the musical metre). Of course, that rhyme class only lasts as long as halfway through the next bar, as illustrated in figure B1.16, making it the fifth bridge rhyme in as many bars.

122 The statement ‘I often drift’ is a cleverly inserted meta-comment upon the poetic and lyrical structure of the verse.
Again, André 3000 emphasises a rhyme in the middle of a syntactic unit by inserting a pause right after the rhyme. Having arrived at a convergent metrical structure (to allow the verse to come full circle), André 3000 rounds off the breakneck passage of high rhyme density and successive bridge rhymes with another of his trademark secondary rhyme inserts (‘be it’ and ‘needed’), sustaining the high rhyme density without really challenging the dominant lineation. In the end, he concludes the verse by letting the lyrical content arrive undisturbed by alternative lineation. The final four bars of the verse display fully convergent metrical structure, with only a simple added one-rhyme as a final technical flourish.

**Summary**

The discussion of this track introduces some key theoretical issues regarding how rappers create dope rhythmic effects at the macrorhythmic zoom level. Ambiguity in lineation—regarding what the dominant lineation is or what is going on with all the weak alternative lineations—is a parameter that rappers manipulate to great aesthetic effect. Some of these techniques are common and formalised, albeit with variations including various one-rhyme techniques and bridge rhyme techniques. The former are examples of heteromorphies among three main types of lineation evidence (syntax, primary rhyme position and musical metre), while the latter is a different sort of heteromorphy—one which mediates between the dominant lineation and rhyme complexes or secondary rhymes.

There is consistent evidence for lines’ inherent lop-sidedness. As in English poetry in general, the beginnings of lines are ‘loose’, while the ends are ‘strict’ (Fabb, 2002, pp. 173-175), and with regard to all of the heteromorphies in this track, it is the potential line endings which are investigated.\(^{123}\) In instances such as added one-rhymes at the beginning of a line, the alternative lineation derives from the link of the rhyme to the previous line, which introduces

---

123 For rap, in particular, this is also noted by Ohriner in his corpus study: ‘phrase endings are more concentrated at beat four compared to the beginning. This means that although the length of phrases in rapping is variable, differences in length are accommodated by starting later, not ending earlier’ (2019b, pp. 109-110)
ambiguity regarding whether the line has ended or not. In the long lines of André 3000, the ambiguity is created by presenting successive potential line endings, not by unclear line beginnings.

The examples also challenge musical metre’s status as a type of strong evidence for lineation. While the conventional metrical structure in rap flows is convergent, with line and bar boundaries coinciding, it does not seem to be the musical metre in itself which functions as strong lineation evidence. Instead, dominant lineation is always marked by either a syntactical boundary or a primary rhyme, and the role of musical metre is to either strengthen or weaken the other two. In an example like ‘Aquemini’, with its extensive divergent metrical structure, musical metre does not play a strong role in determining the dominant lineation until other lineation evidence repeatedly emphasises the metrical boundaries. In fact, the passage with repeated one-rhyming (where the dominant lineation is consistently displaced by one beat position in relation to a convergent metrical structure) is no less metrically stable than the passages of convergent metrical structure. This indicates that repetition, symmetry and parallelism between lines is more significant for creating stability/predictability than strict adherence to the musical metre. Tradition, convention and pre-existing expectations are not enough to empower convergent metrical structure. It must be established in some way by the flow for it to have much referential power. We see this clearly in ‘Aquemini’, and we will see it in several more examples in the next track.
Track B2: Interacting with the *gravity of the form*

Track A6’s overview of the development and formalisation of rap music introduces the concept of the *gravity of the form* as an organising principle for rap flows. Along the lines of the *expected convergence* between lines and bars (convergent metrical structure), the gravity of the form points up a rap flow’s tendency to adhere to the boundaries of the larger symmetrical subsections of the common verse form. Even when the flow embarks on passages with divergent metrical structure, there remains an expectation of convergence of line boundaries and the ends of four- and eight-bar sections. In this track, I will develop this concept further by looking at the techniques which rappers use to emphasise and reinforce the common verse form with (more or less) symmetrical blocks; employ the gravity of the form to punctuate or enhance particular rhythmic or poetic gestures; or even actively work against the order and symmetry of the expected formal structures. These techniques range from the simplest kind of structural emphasis, like placing prominent gestures—for example, a rhythmic gesture which contrasts with the rest of the flow/song or particularly standout rhymes or rhyme techniques—at the boundaries of formal units, to various ways of dividing verses into sections through variations in lineation structure or rhyme. There are also examples of rappers suspending formal gravity to either create a sense of freedom from form (see ‘Aquemini’ in the previous track) or perhaps underpin either a ‘groove mode’ or ‘song mode’ of listening (as discussed in track A6). In the previous track, I presented several techniques which create bridges or connections between larger rhythmic units or rhyme complexes. Variants and combinations of these or similar techniques can also be employed to emphasise or contradict the gravity of the form, as we will see throughout this track.  

**Grand gestures**

One way in which rappers interact with the form is by using *grand gestures*—an umbrella term for those prominent gestures which stand out in the flow and are placed at very specific points within the standard rap verse. They are often accompanied by a change in the musical background to further emphasise their relative ‘importance’ as rhythmic statements. I will begin this track by presenting some examples of grand gestures which emphasise the boundaries of the common formal template of the sixteen-bar rap verse.

---

124 Note that techniques covered in later tracks can also work as gravity-enforcing gestures if they are placed in prominent positions, even if they are not necessarily analysed as such when they are introduced. The possibilities for techniques’ interactions with form are virtually limitless.

**Figure B2.1:** Grand opening gesture of Gatas Parlament, ‘Stem Gatas Parlament’, verse 2, from 1:47. The transcription is a quantised representation of the rhythmic structure. Blue brackets show syntactical boundaries. Green emphasis added to barline to emphasise the musical metre. Weak/hidden rhyme in stippled red outline. Alternative figures for beat 3, bar 2, provided. Expressive timing means there are several acceptable interpretations.

The initial figure of eighth-note triplets represents a clear break from what has happened in the previous verse and the extended scratch break between the verses, where all instances of triple subdivision have been at the sixteenth-note level (the quadruple subdivision which is the primary density referent is also heavily swung—approaching a full 2:1 ‘shuffle’ ratio at times and blurring the boundaries between duple and triple subdivision of the eighth-note subdivisional layer). The heavily syncopated off-beat phrasings in the second halves of both bars are also prominent rhythmic features, and Borgersrud *drags* them heavily as well, so that it is no longer clear what is on and off beat. Interestingly, it would appear that the rhythmic gesture and expressivity have been prioritised over the clarity of the rhyme. The initial plan with the lyrics seems to have involved a compound three-syllable rhyme of ‘er det når’/’fjerde år’, but thanks to the position mismatch and lack of rhythmic rhyme (the rhyming figures are different—syncopated eighth/sixteenth/sixteenth versus three sixteenth notes), one barely registers any rhyming function. This hidden rhyme, which is probably more often experienced as a non-rhyme (that is, as a broken expectation of a rhyme connection between the bars) is in itself a novel choice by Borgersrud, and the resulting lineation is also unclear. The most intuitive interpretation, in my opinion, is that there is one long line spanning the two bars (as indicated by the darker blue in the figure). Syntactically, the text can also be divided in different ways: ‘kan du fortelle meg:’, ‘hva er det når de kaller det demokrati’, and a clause of ‘men kun hvert fjerde år’ could conceivably be considered three different segments, or (more
probably) two segments with the latter two joined together. The convergent version, which is probably the second most intuitive after the long one, would still split up the compound hidden rhyme across the two lines, as ‘når de’ is clearly a pickup to the second bar/second line in a convergent metrical structure. This is yet another reason for the rhyme being hidden—even if the potentially rhyming syllables are placed at the end of the bar, they are not at the end of a line, which weakens our expectation of rhyme. The combination of all these prominent factors—new subdivisional structure, expressive timing, uncommonly long or unclear lineation, and lack of hidden rhyme—adds up to a grand opening gesture which captures the attention of the listener at the very beginning of the verse.

One particularly common way to set the stage for a grand gesture is through a cut, where the musical background is either completely muted or reduced to a minimum, typically by removing drums and other instruments with sharp onsets. The resulting blank canvas ensures that the rap flow receives the listener’s full attention, free of any competing rhythmic onsets. In turn, the rapper can freely introduce rhythmic gestures which deviate from what precedes and follows the cut. Extended cuts are commonly placed at the beginnings or ends of verses, but there is nothing which says that they cannot be placed anywhere within a commonly structured verse. Like the previous example, the passage in figure B2.2 is the (grand) opening of a verse, this time over an extended cut lasting two full bars (except for a kick drum on the initial downbeat and a single snare on the final backbeat).

125 The term ‘cut’ is taken from Snead (1984), who uses it as an umbrella term for different types of cut-like gestures in Black music and literature. I use it for only one of the gestures Snead discusses. To avoid any potential confusion with other applications of the word ‘cut’, I italicise the term when it used to refer to the specific musical application explained here.

126 Amongst the most archetypical examples of this is the final bar of the rap on Michael Jackson’s ‘Black or White’ (1991) by producer and co-writer Bill Bottrell, where the drums stop on beat 1, allowing the punchline ‘I’m not gonna spend my life living a colour’ to take centre stage undisturbed.

127 Some cuts, particularly shorter ones which span only a few beats, might well have been added after the rap was recorded to emphasise particular passages instead of being part of the musical background originally. In my anecdotal experience, this is overwhelmingly the case. When rapping on other artists’ projects, as well, I might request cuts at specific places to coincide with specific parts of a pre-written verse.
Figure B2.2: Grand opening gesture of Big Boi’s verse on OutKast’s ‘Skew It on the Bar-B’, from 1:50. Alternative transcription of late, expressively timed syncopated figure in parenthesis above. Double-planed (one-syllable and two-syllable) rhyme instances below.

Like Borgersrud in figure B2.1, OutKast’s Big Boi opens his grand gesture with a row of staccato syllables which contrasts with the preceding flows (in the verses by André 3000 and Raekwon, respectively). Big Boi’s expressive timing, as well as the lack of any other rhythmic reference, blurs the subdivisional categories slightly, so that the figure on beat 3 in particular seems slightly ‘open’ as to which category its syllables belong to. This is partly due to their duration (which are somewhere between sixteenth and eighth notes) and partly due to the parallelism and semi-rhyme in Big Boi’s comparison of his use of rapping (‘bust raps’) with drug-dealers’ use of guns (‘busts gats’), which invites the listener to experience them as a rhythmic rhyme as well.

In the second bar, Big Boi reintroduces the stable sixteenth-note density referent, and the grandness of this part of the gesture is not in the rhythmic novelty but rather in a clever ‘double-planed’ rhyme structure. The first rhyme connection we experience as line-final is between ‘gats’ in bar 1 and ‘axe’ in bar 2. This one-syllable rhyme is strengthened by the metrical positions (both rhyme instances are syncopated to beat 4) and the expectation of convergent metrical structure, as well as at a thematic level because a ‘gat’ (slang for handgun) and an axe are both weapons. However, Big Boi has a surprise in store for us as he tacks on the words ‘or the hatchet’, yet another weapon, at the end of the second line. This means that rather than a single syllable primary rhyme, there is a strong two-syllable rhyme connection between ‘gats, shit’ and ‘hatchet’, creating a one-rhyme structure somewhere between an added one-rhyme and a simple one-rhyme. The way Big Boi drags the word ‘gats’

---

128 I hesitate to consider ‘bust raps’ and ‘busts gats’ a true rhyme. There are so many weakening factors (the close proximity, the lack of perfect rhythmic rhyme, the combination of identical rhyme on the first syllable and only assonance on the second, the less stressed syllable) that I experience it rather as a speech-sound parallelism more akin to alliteration.
in bar 1 then becomes part of making ‘gats, shit’ and ‘hatchet’ rhyme more rhythmically as close to a pair of eighth notes in both instances (he is even slightly early on the initial ‘sh-’ sound on the word ‘shit’).

Grand gestures need not be a technical flourish or a ‘flipping of the flow’, of course. The dramatic use of pauses or a musical meta-comment or conversation between the flow and the beat are other ways of making a grand musical statement. Take, for example, the halfway point of the second verse of Nicki Minaj’s ‘Starships’ (2012, from 01:35), where the seventh bar sets up with Minaj rapping the first phrase of the topical traditional lullaby ‘Twinkle, Twinkle Little Star’, after which she then goes silent in bar 8, allowing the musical background to respond by playing the melody of the lullaby’s second phrase. Another technique involves winding it down at the end of the verse, by breaking out of the rhythmic regularity, changing from rapping to speech, and/or breaking structural expectations by omitting rhyme, for example. An example of this can be found by returning to Big Boi and ‘Ms. Jackson’ and the end of his second verse (the third of the track), where he forcefully tacks on ‘you and your mama!’ after his final rhyme complex (listen from 3:09 to hear the context). There are a multitude of variants of this type of grand gesture, and they are very often found at the end of verses or (as in ‘Starships’) at the end of large structural blocks. To make gestures like this particularly grand, rappers might build up to them by cadencing.

**Cadencing**

In colloquial rap discourse, the word ‘cadence’, like ‘flow’, is used rather indiscriminately to refer to a wide variety of things; it is also defined in many ways which are rarely consistent or especially specific. At times, it appears to be a synonym for flow; other times, it recalls the way Edwards (2009) describes ‘delivery’, particularly in relation to the use of pitch. Thus, one could perhaps compare the concept of cadence in a rap flow with the phonological (melodic) concept of boundary tones, which again relates closely to lineation. Here, I will attempt to use the term as it is understood in traditional music theory—as a sequence indicating finality. This relates to the gravity of the form in several different ways. First of all, when rappers *cadence* (as a verb), the degree of finality relies upon the manner in which various types of evidence for lineation coincide, and an expected coincidence between line boundaries and the boundaries of symmetrical sub-sections of a verse is particularly salient. Secondly, rappers

---

129 Similar short ‘tack-ons’ where the rapper winds down the flow by adding a non-rhyming statement after the final rhyme can be found at the end of Eminem’s ‘Stan’ (2000, from 6:33) and the final verse (by Big KRIT) of A$AP Rocky’s ‘1Train’ (2013, from 5:51), for example.
will also strengthen the relative impact of a cadence by withholding coincidences of lineation and metrical boundaries. This is a means of increasing tension by delaying release which is again analogous to functional harmony.

While the concept of cadences and their structural and formal weight and gravity are all well understood from the tradition of harmonic analysis in music, there are other interesting approaches to studying cadences in poetry. A prime example is the work of Winfried Menninghaus and colleagues wherein physiological measures of ‘being moved’, such as goosebumps and chills, are found to correlate with cadentially heavy positions within poems:

A study on the physiology and neural correlates of reading emotionally moving poems revealed that peak moments of states of being moved—as marked by chills and goosebumps—are sensitive to an important compositional feature: They typically occurred toward the ends of lines and stanzas, with the intensity increasing the more the poem approached its closure in the final line. Hence, feelings of being moved appear also to be sensitive to the music-analogous tension–resolution structure of poems: With each additional line, the predictions of readers regarding the poem’s formal patterning and its overall trajectory, including its content, become increasingly strong, and so do the rewards of the resolution perceived both at intermediate closing points of the composition (cadences of lines and stanzas) and at the poem’s final conclusion. (Menninghaus et al., 2019, p. 177)

Both the emergent form of a piece of art and pre-existing notions of structure and form are clearly important virtual reference structures to which our physiological response system adheres. It seems clear that we, as listeners, thoroughly enjoy the relative predictability of cadences towards specific formal boundaries. Returning to the interaction between prediction and ‘top-down’ or ‘bottom-up’ driven attention discussed in track A3, one could theorise that there is a stronger expectation of adherence to the boundaries of larger symmetrical blocks than to those of smaller metrical units (bars). It is probable, then, that this expectation being fulfilled or broken is a relatively more marked event than most—one which leads to an involuntary bottom-up driven response of ‘catching our attention’ even if we are not consciously attending to the symmetry of the formal sections or subsections. It seems clear that there are some expectations whose fulfilment is more enjoyable than others, and artists (whether they create written poetry or rap) can take advantage of this by increasing the relative tension to be resolved at these strategic points. To make clear the difference between techniques which actively exploit these structural expectations and the simple coincidence of
lineation evidence and the boundaries of larger structural blocks, I reserve the term _cadencing_ for instances where there is some sort of added structural tension from divergent or ambiguous metrical structure. Not all resolutions are the result of cadencing, but all cadencing leads towards some sort of resolution.

An archetypical way of cadencing is to use divergent metrical structure to delay a coincidence of line ending and bar ending until the final beat of a hypermetric unit. Figure B2.3 shows an example of a hypermetric unit of four bars displaying a divergent metrical structure before finally cadencing and ‘landing’ on the final beat 4.

![Figure B2.3: Divergent metrical structure and cadencing in the first hypermetric unit of MF Doom’s ‘THAT’S THAT’ (2009), from 0:00. Black dots indicate beat positions, coloured bars show lineation structure (with the colour indicating rhyme complexes).](image)

This excerpt from ‘THAT’S THAT’ is interesting in that the lineation is not particularly clear. In fact, there are several potential lineations, but few of them are supported by strong evidence. The syntax is unclear and ambiguous, and the rhyme density borders on full.
saturation. Thus, changes in primary rhyme—that is, the division into rhyme complexes—represent perhaps one of the stronger types of lineation evidence in the passage. The line of short /i:/ rhymes (marked in green) is attached through proximity and delivery to the following /ɔ/ rhymes, and MF Doom only audibly pauses and breathes at the end of line 2 (after ‘mirror broke’). This pause becomes an argument for dividing this section into two groups of nine and seven beats, respectively, subdivided as indicated in the figure. No matter which way one interprets the dominant lineation, there is no convergence of line and bar endings until the end of the hypermetric unit, and this delayed rhythmic ‘resolution’ is an example of a rapper cadencing using metrical structure.

The resolution of a cadence does not necessarily need to happen on the final beat of a formal unit; it is also possible to build up structural tension and resolve it at another metrical position, as we can see in figure B2.4.

Figure B2.4: Cadencing through increased rhyme density in Side Brok’s ‘Heimegut’ (2013), from 0:27. Transcription from Oddekalv (2017).

The first two bars of the figure display convergent metrical structure and introduce the three-syllable primary rhyme (via an identical rhyme, in the first instance) which in practice spans two beats thanks to a striking syncopation on the final syllable. In the third and fourth bar, the rhyme complex continues, while the rhyme density increases. The rhyme falls first on the third (and syncopated fourth) beat of the third bar, then the first (and syncopated second) beat of the fourth bar, and this increasing rhyme density functions as an accelerando cadencing the line towards a resolution on beat 1 of an extra added bar (a one-rhyme). Rapper Runar Gudnason130 mimics the increased tension of the rhythmic structure in his delivery, rapping

---

130 Gudnason performs under several different alter egos in Side Brok. Sjef R (the character he inhabits in the example analysed here) and Thorstein Hyl III are differentiated through on-stage garb and a clearly different vocal delivery (Thorstein Hyl III is characterised by a higher pitched and more nasal voice). The character Sirkel Sag belongs to the same universe but is a solo act in addition to his occasional cameo in Side Brok. Visually, Sirkel Sag is portrayed by another Side Brok member, Bård ‘Skatebård’ Aasen Lødemel, but the raps are written (and, I suspect, also recorded) by Gudnason.
with more and more pathos as the rhyme density increases before initiating a decrescendo towards the final resolution. This example highlights how the first downbeat after a symmetrical block (or expected formal subsection, one should say, because in this case there is a single extra bar tacked on at the end of a verse) is experienced as metrically significant. This is the norm in most music, and rap music’s focus on beat 4 as the common resolution position is atypical. But here, Gudnason shows that a cadence towards a resolution on beat 1 is an option in a rap flow as well. The final word tacked on at the end—‘Nei’ (meaning ‘no’)—functions both as a grand gesture of sorts—a comment on both the flow and the content preceding it—and as an entrance cue setting up the start of the next verse. It both answers the question set up by the cadencing phrase and takes us outside of the flow for a split second because it breaks all structural expectations regarding a segment of flow—it neither rhymes nor is ‘a line’.

‘Helt om natten, helt om dagen’: Lars Vaular’s mastery of gravity
A verse that employs nearly every single technique covered in this track is Norwegian rapper Lars Vaular’s first verse on the title track of his 2010 album, ‘Helt om natten, helt om dagen’, visualised in figure B2.5.
The first thing that stands out in figure B2.5 is the way in which Vaular structures the verse as four distinct blocks of four bars. The first parameter he manipulates to accomplish this involves employing symmetrical rhyme complexes that span either convergent line pairs (two bars) or full blocks of four bars. The figure also shows that he plays with the predictability of this approach throughout the verse; rather than going technique-by-technique, this analysis will follow the verse as it unfolds through time.
In the first four-bar block Vaular establishes a familiar convergent metrical structure wherein primary rhymes and syntax coincide with the boundaries of the musical metre. Each line is also symmetrically subdivided in half through syntactical units as well as a repeated pattern in each line pair wherein the set-up-line establishes the primary rhyme by repeating it (twice in line 1, thrice in line 3) and the second line features only one instance of the primary rhyme falling squarely on beat 4. The dominant lineation is never in doubt, but the alternative half-line lineation is also weakly present.

Having thus established that both he and the listener share an awareness of rap’s most common metrical structure, Vaular then departs on a four-bar, fully divergent cadence built around a single two-syllable primary rhyme which he reinforces in a very particular manner. Figure B2.6 indicates Vaular’s complete consistency regarding the subdivisional durations of the syllables—every rhyming syllable is an eighth note, which introduces a rhythmic rhyme of two consecutive eighth notes, and every non-rhyming syllable is a sixteenth note. This way, rhymes are strengthened not only by the stacking of parallelisms but also by eliminating any coincidental rhythmic similarities between rhyme and non-rhyme.

**Figure B2.6:** Transcription of bars 5–8 of ‘Helt om natten, helt om dagen’, from Oddekalv (2017). Consistent difference in note duration on rhyming (red) and non-rhyming (blue) syllables.

The lineation in this hypermetric unit is less clear than in the preceding one, and Vaular places the first stressed syllable of the primary rhymes alternately on and off the beat (first instance on, then two off, then on again, apparently to avoid creating a predictable pattern). Most importantly, he also eludes any convergence between line endings and bar endings up until the final downbeat of the four-bar block. The two times when primary rhyme instances are placed around beat 4 (the two first bars of the period) are clearly not line-final rhymes: the syntax places them at the beginning of their respective lines, and the stressed first syllable of each rhyme instance is off beat, slightly weakening their connection to any ‘line-endingness’

172
otherwise invoked by their musical metrical position. If anything, this four-bar block is an indication that musical metre in itself is a rather poor type of lineation evidence, as one is never tempted to interpret line and bar endings as coinciding in the first three bars. The final bar, on the other hand, is clearly the culmination of the three preceding bars of cadencing, and the final primary rhyme instance of ‘peace out!’ lands squarely on beat 4, indicating the end of the symmetrical formal subsection while marking the halfway point of the sixteen-bar verse.

Having performed such a striking gesture of finality at the end of the second four-bar block, Vaular also cannot resist the temptation of playing with listener expectations at the beginning of the second half of the verse. He starts out the first line of the new formal section by continuing to rhyme with the preceding primary rhyme—that is, the one we thought we were done with! After a short caesura, however, Vaular playfully tacks on another clause to the line, introducing what will become the new primary rhyme. This line (or these lines) displays true ambiguity in its lineation, at least as the music unfolds. First, the line is experienced as closed after ‘de sa’—that is, as a continuation of the patterns from the previous block (same primary rhyme and a segment spanning three beats)—an interpretation supported by the pause Vaular inserts here. It is as though he wants us to think that he is not yet finished with the /ɪ*ɑ/-rhyme or the strict separation of rhyming eighth notes and non-rhyming sixteenth notes. However, with the reveal of the ‘. . . eller ka de sier’ clause, we are forced to reinterpret the dominant lineation. This short phrase (which translates to ‘. . . or what they say’) cannot really stand on its own as an independent line, and this makes an interpretation of a long line—following the syntax—as the dominant lineation into a plausible interpretation. The situation remains ambiguous, however, because even when the new primary rhyme is confirmed (by the short phrase ‘ka politiet sier’), the imbalance in line length between the two parts of the line pair (spanning around six and two beats, respectively) might lead us to interpret the first line as two short lines (of around three beats each), even if the syntax disagrees.

The reason I prefer the interpretation of a long line as the dominant lineation, with the division into two after ‘de sa’ as a weakly present alternative lineation, involves the way in which Vaular has used the gravity of the form. While musical metre in itself is not strong lineation evidence, Vaular’s invocation of a sense of finality with the cadence towards the final beat of the first half of the verse makes me hesitant to view the first part of the first line in the second half of the verse as ‘part of’ the passage which precedes it. The expansion of the rhyme complex, coupled with the bridge between the symmetrical blocks created by the ‘de
sa’ bridge rhyme, is an elegant means of linking two separate units, but they remain clearly separate nevertheless. The symmetrical (4+4+4+4) structure and the way in which Vaular reinforces it stand up to the challenge of the bridge rhyme, and that could be why it stands out as such an elegant compositional feature.

After the twist of the first line, the remainder of the third four-bar block re-establishes a more predictable convergent metrical structure, where the two stressed syllables of the three-syllable primary rhyme are placed on the third and syncopating to the fourth beat. Again, the rhyme complex coincides with the symmetrical structure, perpetuating the reinforcement of the four-bar blocks as the most important structuring idea of the verse. The only additional flourish Vaular adds here involves the secondary rhymes in the third line, where ‘dies’ and ‘beviser’ rhyme partially with the primary rhyme, creating an increased rhyme density without saturating the line with primary rhyme instances (as was the case in the second four-bar block).

The final block is, in a way, a nod to the opening of the verse, in that it is split into two with a change in the primary rhyme rather than reliant upon one primary rhyme throughout the block. The first line pair also features a symmetrical half-line phrasing, with primary rhymes on the beat 2s in addition to the beat 4s. The final two bars, on the other hand, are much less conventional. Instead of offering a clear lineation with rhymes marking the ends of the lines, Vaular employs a sort of ‘double-rhyme’ strategy, with two weaker rhymes competing for the role of primary rhyme while crossing over one another. This makes for a peculiar passage, as it feels like something rhymes but it also feels like the final line ending does not rhyme (as one would expect), as can be seen in figure B2.7.

![Figure B2.7: 'Hidden' crossing rhymes in the final bars of 'Helt om natten, helt om dagen', verse 1. Brackets indicate syntactical units; the two competing rhyme classes are indicated below. Notice that Vaular pronounces](image)

131 Except for the bridge rhyme at the beginning of the first line of the block, of course. However, the rhyme instance of the bridge rhyme does not function as a primary rhyme when the line is established, so it can be considered a secondary rhyme in this new section. Rhyme complexes have unclear beginnings and clear(er) endings, of course, so determining where a rhyme complex begins is both challenging and relatively insignificant to a flow’s structure compared to where it ends.
the penultimate syllable ‘-ner’ with an /ə/ sound (as is typical of his dialect), so it is phonologically parallel to the ‘-da’ syllable in ‘agenda’.

The figure indicates how the two competing rhyme classes coincide with the syntactical units and the musical metre, respectively. Due to the order in which the rhyme instances are revealed, the first rhyme class takes precedence. It is quite weak—only four vowel sounds which do not even match perfectly (Vaular pronounces the final unstressed syllables ‘-ne’ and ‘-er’, with slightly different ‘e’-sounds)—but since these rhymes coincide with the ends of syntactical units, we are predisposed to think of them as rhyming. The other rhyme class is placed at the expected beat 4 position of both bars, but due to the competition from the other rhyme class (which is strengthened by its coincidence with the syntax), this second rhyme class might not be heard as a rhyme at all. After the second syntactical unit ends—establishing the first primary rhyme—we tend to no longer consider anything from before that line boundary as a valid first instance of a primary rhyme. In short, if we experience the first rhyme class at all, the second rhyme class’s impact is greatly diminished. It might be that if the end syllables of the first two short syntactical units clearly did not rhyme, we would hear the speech-sound parallelism between the syllables at the ends of the bars. Also, if Vaular were to insert a clear pause in the second syntactical unit, spelling out through his delivery the fact that ‘agenda bak’ is the end of something, we might also be primed to listen for something rhyming with it. However, he does not, and there is no doubt that the dominant lineation persists through ‘det eg sier’. In the end, one certainly does not expect the final /e:*a:/ rhyme, as it is not reminiscent of any previous line endings or established rhyme classes and comes across as an unexpected non-rhyme or a feeling of being unsure of whether one missed a rhyme somewhere.

‘F*ck Your Ethnicity’: working against symmetry

Verses like the one in ‘Helt om natten, helt om dagen’ exemplify the fact that rappers are proficient at diverging from and challenging the metrical authority of convergent metrical structure but will still adhere to a fully symmetrical and predictable organisation of subsections within a verse. This combination is, in some ways, both radical and conventional at the same time. On the one hand, the formalisation of the common verse template allows rappers a greater degree of freedom as to what they can do within its boundaries; on the other hand, it provides a sense of predictability and structure that allows for seamless integration of different elements within a verse.

---

132 This is yet further proof of the need for triangulation among several parameters when establishing lineation, and of the fact that musical metre does not imply lineation in and of itself. Since there is no syntactical unit boundary and no pre-established rhyme or exaggerated pause, the musical metre does not impart any structural significance to the syllables on/around beat 4.
hand, adherence to its symmetry invites one mode of listening and discourages another. Recall from track A6 that a predictable formal structure which underpins and reinforces the listener’s expectation of the arrival of a subsequent ‘different’ formal unit invites a *song mode of listening*. This goal-oriented mode contrasts with a mode of listening wherein one immerses oneself in the music’s continuous forward movement without being oriented towards a resolution or formal break—that is, a *groove mode of listening*. There are many features present in a hip-hop track which invite both modes of listening—the (often) short repetitive basic units which build up the beat represent a typical groove mode—inducing feature, but these basic units are usually pre-organised into fully symmetrical blocks, which increases their formal predictability and suggests a song mode of listening. The flow is in a unique position to potentially undermine symmetry, though it certainly does not have to do so. Asymmetrical lineation can obfuscate the cumulative nature of symmetrical blocks, wherein groupings of four bars are combined into blocks of eight and then sixteen, presenting predictable locations for the formal breaks which define the song mode. The example used in track A6—Warren G and Nate Dogg’s ‘Regulate’—instigates a groove mode of listening through its sheer repetition of structurally similar four- and eight-bar blocks which goes far beyond the basic sixteen-bar verse framework. If ‘Regulate’ had a formal break after sixteen bars rather than continuing on for a full thirty-two, it might not have managed this effect. In shorter verses, particularly those that conform to the genre-typical verse template of sixteen bars (which in itself is an invitation to a song mode of listening), the flow must take a different approach if it is to keep the listener moving within it (in a groove mode of listening) rather than with it towards an expected destination. The musical background is amenable to being pulled in either direction. If the flow emphasises a predictable organisation of the beat’s basic unit(s), it might accommodate a song mode of listening. Conversely, if the flow destabilises the symmetrical, additive organisation of basic units, their basic groovy repetition will be foregrounded instead.

The flow’s ability to create a continuous forward motion which destabilises any anticipated structural symmetry has been touched upon in previous tracks including track B1, where the long and unpredictable lines of André 3000’s flow on ‘Aquemini’ evokes the liquid fluidity of *flow’s* etymology. The way in which a vocal track (‘the lead’, if you will) can create a ‘dissolution of the hierarchy of sequences’ is also discussed by Danielsen (2006, pp. 174-179) in her analysis of an example of an archetypal ‘groove structure’ track—James Brown’s ‘Get Up (I Feel Like Being a) Sex Machine’ (1970):
‘Sex Machine’, for example, is marked by extraordinary manipulation with regard to its periodic structure. We are almost spun into the song’s rhythmic motion by a melody pretending to be the start of a regular verse. We are tangled in an emerging linearity that then vanishes before reaching closure. (Danielsen, 2006, p. 174)

‘Aquemini’ could be said to exemplify a departure from and subsequent return to a song mode of listening, where the linearity evoked by its initial symmetry and convergent metrical structure gives way to unpredictability and an obfuscation of the song mode’s ‘goal’ before ultimately reintroducing symmetry and a clear(er) expectation of a formal break. The next verse I will analyse—the first in Kendrick Lamar’s ‘F*ck Your Ethnicity’—simply departs from the song mode without looking back.

In figure B2.8, I have attempted to capture some of the most significant parameters of the structural organisation of the first verse of ‘F*ck Your Ethnicity’ using a slightly different method of visualisation than previously. Rather than representing the lineation as clearly as possible, the figure shows the lyrics and rhyme complexes as they line up with the musical metre. This is done to emphasise the relationship between the main dominant metre and the double-time (or ‘half-line’); the lack of symmetrical groupings of bars; and the lack of rhyme at certain expected positions and even over extended sections. This method is also pragmatic, as there is a lot of information packed into one figure (or on one verse, rather), and the hand-drawn lines-versus-bars type of illustration employed elsewhere would be unreasonably cluttered and unintelligible.
<table>
<thead>
<tr>
<th>Bar</th>
<th>Lyrical Text</th>
<th>Musical-Metrical Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fire burnin' inside <strong>my eyes</strong>, this is the music that saved my life</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Y'all be callin' it hip-hop, I be callin' it hypnotize</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Yeah, hypnotize, trapped my body, but freed my mind</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>What the fuck is you fightin' for? Ain't nobody gon' win that war</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>My details be retail, man, I got so much in store</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Racism is still alive, yellow tape and colored lines</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>It's so diverse, They getting off work and they wanna see Kendrick, Everybody can't</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>drive Benz's, and I've been there, so I make it my business to give 'em my full at-</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-tention, ten-hut! Man, I gotta get my wind up. 'cause I got my sins up</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Ooh. Matter of fact, don't mistake me for no fucking rapper</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>They sit backstage and hide behind the fucking cameras</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>I moshpit had a microphone and I tossed it</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Had a brain, then I lost it I'm out of my mind, so don't you</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>mind how much the cost is. Penny for my thoughts. Everybody,</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>please hold up your wallets. Yeah, man, I'm the</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>mailman, can't you tell, man? Going postal</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>never freeze up when I approach you That's starstruck, and roast you (oh my)</td>
<td></td>
</tr>
</tbody>
</table>

Figure B2.8: Lyrical, rhyme-complex and musical-metrical structure in Kendrick Lamar’s ‘F*ck Your Ethnicity’, verse 1, from 1:03. Bar numbers on the left. Colours indicating rhyme classes, with lack of colouring indicating that there is no rhyme in the two-bar cell. Thicker lines demarcate rhyme complexes.

This verse and track breaks with many of hip-hop’s structural conventions. The figure is organised into nineteen lines according to what I consider the dominant musical metre—a slow 75.5 bpm 4/4 emphasised by a booming sampled drum track. I have also subdivided each line into two halves, however, because Lamar’s flow, in conjunction with the musical background, turns this musical metrical structure around by employing a double-time feel at certain points, and (later) by displacing the lines from spanning bars to spanning half-line to half-line. Also, the relatively slow tempo indicates a density referent of thirty-second notes, which is much less common than the sixteenth notes which would normally be the density referent of such a double-time time signature. As we will see in the following analysis, Lamar expertly manipulates this tension between the basic metre and the double-time passages.

Initially, as in many of the previous examples on this side of the thesis, this flow conforms to the expected convergent metrical structure. However, from the very first rhyme complex, the typical duple/quadruple-bar grouping symmetry is broken, as there are three lines/bars in the first rhyme complex rather than two or four. Already at this point, as well, Lamar is

---

As noted in track A6, fn. 88, page 107, this is a feature of some rap flows, particularly when there are competing alternative tempos (that is, when there is a level of ambiguity as to whether the double-time or the half-time of the metre is the ‘actual metre’).
emphasis on the middle of the bars as an alternative organisational boundary (or the ‘half-line’, as derived from convergent metrical structure and discussed in track B1), with most lines in the first part of the verse including rhymes on beat 2 as well as beat 4. Still, this is clearly a weak alternative lineation and not a real threat to the dominant one, as can be seen in line 2 (with no rhyme) and line 5 (with a different secondary rhyme on beats 1 and 2, highlighted in green in the figure). When the sixth line turns out not to be part of a rhyming pair, the balance between the slow metre and the double-time seems to shift slightly. Yet, since the sixth line rhymes with the first three (and these lines are not yet distant enough for that rhyme connection to have disappeared), we are able to cling to the convergent metrical structure and construct a (slightly unstable) song structure—like section spanning six bars.

From here, though, the song mode of listening is directly challenged by Lamar’s flow. He pauses on the first two beats of the seventh bar, starts the next phrase after beat 3 at a breakneck pace, and tumbles into a passage with a divergent metrical structure whose rhymes are scattered throughout the bars.\(^{134}\) It takes a while for a new primary rhyme to establish itself. The first syntactical unit ends with ‘line’, which potentially links it to the sixth bar/line through an identical rhyme, but I do not experience this as a strong rhyme when it appears. This might well be because the preceding section of the flow has been ‘concluded’, and I am awaiting a new primary rhyme in this new passage. The ‘work’/’verse’ rhyme in bar 8 (marked with a teal highlight in figure B2.8) is potentially presented as that new rhyme, but because the line continues and fails to actually end on a rhyme, my intuitive interpretation is that ‘work’/’verse’ is a secondary rhyme and I am still awaiting a new primary rhyme (now to match the line ending ‘Kendrick’). Lamar delivers this but initially fairly weakly, because ‘Benz’s’ not only barely mirrors the vocal sounds in ‘Kendrick’ but also has a very contrasting consonant ending. The rhythmic rhyme and the additional emphasis of Lamar’s delivery make the rhyme connection very clear, however. The passage then continues to reveal a series of similarly weak rhymes (‘been there’, ‘business’, ‘tension’), connected as much through the repeated rhythmic figure and Lamar’s delivery as through their phonological content.

This weak chain rhyme then culminates with a new ‘unprepared’ line ending when Lamar punctuates the passage by exclaiming ‘ten hut!’ and inserts a pause in the middle of bar 10. At

---

\(^{134}\) The new passage is also marked as ‘different’ through a gradual crossfade in the vocal tracks, where the backtrack/dub tracks gradually overwhelm the main track. This production technique mirrors the flow’s emphasis upon the double-time metre and gradually falls back into the conventional balance of a loud main track and quiet backtracks as the breakneck pace subsides and the slow metre re-emerges.
this point, we have been desensitised to what is beat 1 and what is beat 3 (unless we have been directing our attention to certain specific aspects of the musical background throughout the track), but Lamar allows us a hint of convergent metrical structure by using the same half-line phrasing as that in the verse opening (‘with god’ in line 11 does not rhyme with its surrounding parallels, but it is the same rhythmic figure—that is, it rhymes rhythmically but not phonologically). Before we can fall back into a song mode of listening, though, Lamar again disorients us by embarking on a two-bar passage without a single rhyme in it.

Thus far, that is, the dominant lineation has been lining up with the (slow) musical metre—albeit heavily challenged by weak alternative lineations and lacking any symmetrical grouping above the bar level—but in the end Lamar turns the metrical orientation inside out, transporting the line boundaries from beats 1 to 4 to 3 to 2. He does so by continuing to use short rhyming phrases spanning two beats, which means that the half-line phrasing temporarily takes over as the dominant lineation. This shift also coincides with a change in the musical background as the piano chords abandon the two-bar loop of /Am Em/F/ that has grounded the hypermetric symmetry thus far to a more ambiguous sequence. In bar 14, the F major chord of the second bar of the loop becomes the first bar in a new three-bar progression of /F F/Am D/G Dm/ and the piano hammers chords on every 1 and 3. After three such two-beat rhyme phrases, Lamar returns to lines spanning four beats, halving the rhyme density and ending lines on beat 2 instead. He retains this structure throughout the final bars of the verse, only ‘turning it around again’ at the very last moment via another half-line-phrasing passage of three rhyme phrases spanning two beats which rounds off the verse with another six-beat rhyme complex. The verse then ends unexpectedly with a fully conventional end-rhyme-on-beat-4 ending after a very unconventional nineteen bars which did not give the song mode of listening any predictable symmetry with which to work. Lamar signs off with a playful aside which nods to his toying with our structural expectations: ‘Oh my!’.

Summary

Rappers have many different ways of manoeuvring within the established structural norms of the common verse template by either emphasising, destabilising or otherwise exploiting the gravity of the form. At the boundaries of the main symmetrical blocks, a grand gesture may

---

135 The synth motive that accompanies the A minor chord from the first loop continues to appear on the second loop’s A minor, giving the second bar of the new progression an extra emphasis which both increases the new progression’s ambiguity and strengthens the feeling (or rather illusion) of hypermetric regularity for the final two bars—that is, lining them up with the reintroduction of the original two-bar loop starting on A minor when the chorus arrives.
be used to capture the listener’s attention or finish off a verse or section in style. These stand-out flourishes (or even stand-out subtle details, however oxymoronic the idea might seem) will often coincide with a cut, where all or part of the musical background will be removed to let the gesture appear uncontested.

Rappers build up rhythmic tension prior to an expected release at the ends of verses or hypermetric blocks by cadencing—for example, by using divergent metrical structure throughout a section before the final line ends as anticipated on beat 4 (as exemplified by MF DOOM and Lars Vaular). Another approach was unpacked in the example by Runar Gudnason, where a gradual increase in rhyme density builds up tension in a cadence which ends with a one-rhyme.

Verses by Lars Vaular and Kendrick Lamar showcased two different approaches to subdividing a verse. The former emphasised and played along with the gravity of the form, while the latter resisted expectations of symmetry. Vaular actively uses the gravitational force of a symmetrical division of sixteen bars into four blocks of four to reinforce cadencing, play with our expectations of closure using a bridge rhyme, and ‘justify’ a hidden rhyme. Lamar, on the other hand, first disposes of the hypermetric symmetry, then turns the metrical structure ‘inside out’ through a manner of phrasing which casts doubt on whether the dominant metrical structure spans from beat 3 to beat 2 or beat 1 to beat 4.
Track B3: When timings meet: Microrhythmic techniques

In the previous two tracks, we have looked at techniques rappers use to create rhythmic tension and ambiguities via the manipulation of expected convergence at the macrorhythmic zoom level. The divergent parameters have involved larger rhythmic building blocks like musical and poetic metrical spans—lines and bars, various types of lineation evidence, and rhyme complexes and symmetrical blocks. In this track, however, we will zoom in on the microrhythmic zoom level, a realm of expressive timing, analogue features (that is, continuous within-category modulation) of canonical digital time-value classes (recalling Kvitte’s digital/analogue dichotomy from track A3), and the systematic and unsystematic variation of the temporal positions of rhythmic events. These are all different ways of approaching similar experiential phenomena. Instead of being the result of clear categorical divergence and convergence, the techniques described in this track are created by the manipulation of slight (and not-so-slight) mismatches between categorical rhythmic information in two or more layers of the composite auditory stream. Having committed to focusing primarily on the largest and smallest rhythmic units, I will not linger upon the mesorhythmic or quantised zoom level as such. However, the analysis of microrhythm and expressive timing does engage with the ways in which expressive microrhythmic analogue features influence the digital ‘quantised’ rhythmic structure. Thus, some quantised representation will be central to most of the analyses, even though the analyses themselves are focused on microrhythm. The techniques showcased in this track range from those which are a result of ‘straightforward’ variation within clearly defined time-value classes (that is, the analogue modulation of the digital) to truly ambiguous rhythms which toy with our initial interpretations and subsequent reinterpretations of categorical rhythmic information when we follow them as they unfold. In reference to the discussion in track A6, this track explores the intersection of expression and structure, and in particular how the expressive informs or dictates the structural.

Metrical anchors: ‘Waterfalls’

The vocal track is the soloist of the rap track’s orchestra, and it tends to be the layer(s) of the composite auditory stream to which we attend most closely when listening. And while we have already established that stressed syllables do not have to coincide with strong positions in the musical metrical framework (they might, for example, syncopate or create cross-rhythms instead), there is an expected convergence at play. Or, put differently, all syllables boast several categorical belongings, including their specific position within the metrical latticework (for example, ‘the final sixteenth note of the second beat’) and the binary category
of on or off beat. When there are conflicting clues in different layers of a stream as to where the boundaries of certain rhythmic categories are—that is, when there are multiple rhythmic events from different layers which communicate the same general category (like a beat or a sub-beat), but the events do not share the exact same temporal position—stressed syllables have a crucial structuring role. We expect convergence of stressed syllables and main rhythmic categories, meaning that when the boundaries are unclear, a stressed syllable will be definitive for the category. Stressed syllables function as metrical anchors, anchoring the unfolding rhythm to our music-metrical reference structure, or, more precisely, anchoring our music-metrical reference structure to the unfolding rhythm.

An example of a metrical anchor can be found in Lisa ‘Left-Eye’ Lopez’s verse on TLC’s international hit track ‘Waterfalls’ (1994). Throughout the verse, Left-Eye keeps the timing of the flow quite close to the rhythmic categories indicated by the musical background, with heavily swung sixteenth notes (around a 2:1 ratio), then departs on a passage with much more expressive timing, as illustrated in figure B3.1.

**Figure B3.1**: Metrical anchor in Lisa ‘Left-Eye’ Lopez’s verse on ‘Waterfalls’ (TLC), from 3:31. Red dots on middle grid indicate the temporal position of syllables compared to a (straight) sixteenth-note framework derived from the musical background. Transcription above is what I consider the most probable interpretation of the unfolding rhythm. Transcription below is (an improbable) one derived from a close attending to categorical information derived from the musical background.

Figure B3.1 is an illustration of a potential listening experience (derived from my own) using a metrical anchor to structure the rhythmic interpretation. As the flow abandons the previously established swung-sixteenth-note category, there ensues a passage of unclear
rhythmic categorisation which forces the listener to await a metrically defining event with which to make sense of it all. This event—the metrical anchor—appears in the second syllable of ‘enduring’ and defines the one (the downbeat) with gusto, strongly emphasised by Left-Eye’s delivery (and particularly the pitch contour).

While it is evident from the figure that the syllable does not coincide with the ‘objective’ temporal position of the downbeat indicated by the musical background, I would argue that most listeners will experience the ‘dur-’ syllable as the ‘one’ rather than as a displaced stressed syllable on the second sixteenth-note position of the beat. There are many reasons for this. We have already established the obfuscation of the subdivisional categories in the passage preceding the metrical anchor (the first four syllables share a very similar duration, but not one which fits any sensible subdivisional category). Additionally, there is no evidence for the ‘displaced’ interpretation other than the syllable’s temporal position. It does not create any type of crossrhythm or counterrhythm by cooperating with other displaced heavy syllables, nor has the rest of the verse hinted at this type of rhythmic displacement as a recurring technique (generally, Left-Eye’s displaced syllables are syncopations anticipating beats throughout the verse). The distance between the ‘expected’ position of a downbeat event and the metrical anchor is also small enough to be within what a listener in this context might find is acceptable expressive timing. The enormous difference in prosodic emphasis between the preceding weak syllable and the stressed metrical anchor further emphasises the relative importance of the metrical anchor as the structuring rhythmic event. In a hypothetical variant of the flow where Left-Eye adds prosodic emphasis (through pitch, amplitude or even elongation) to the weak ‘en-’ syllable, one might well have interpreted the on- and off-beat categorical belonging of both it and the metrical anchor differently, but there is nothing in Left-Eye’s delivery which indicates anything other than the metrical anchor being allowed to define the downbeat. The musical notation in figure B3.1 illustrates the difference between a ‘measured’ transcription according to a specific ruler (in this case, strict isochronous subdivision categories lined up with the musical background) and a transcription which attempts to capture a representation of a listener’s interpreted rhythmic figure(s). Hopefully, it should be evident that the bottom ‘measured’ transcription is nonsensical when it comes to

---

136 Not necessarily ‘expected’ from our listening experience but rather from the analytical framework of using the specific ruler of a relatively quantised musical background.
137 I have done side-by-side analyses of this passage using both the original studio recording and other performances of the verse which I have found online, and the passage is consistently ‘expressive’ without in fact being consistent as to the ways in which this expressive timing varies in relation to the musical background. There is nothing other than the metrical anchor defining the downbeat in any of the alternative versions either.
representing any listener’s immediate interpretation of the rhythm as it is unfolding. We are, for example, probably not capable of identifying or categorising the first four notes as having a note value of a sixteenth note plus a sixty-fourth note (which the ‘+’ notation attempts to indicate). Importantly, the top transcription is probably not representative of a listening experience either, because we most likely structure our rhythmic interpretations around the most salient and clearest rhythmic categories. The events which are marked and stand out will anchor our figures just as heavy syllables anchor our metrical understanding. The less salient events surrounding the marked ones might be processed more approximately or grouped together with a quite liberal use of ‘cognitive quantisation’—that is, one does not necessarily sort each rhythmic event into the ‘most correct’ category but rather into whichever category ‘works’. Any one specific transcription or explanation of an expressively timed passage such as this is necessarily an abstraction of sorts. Who is to say that interpreting ‘such an en-’ as a sixteenth-note triplet (whether or not one is conscious of the concept) is in fact part of how the listener experiences this musical passage? We might group certain syllables together when they span a beat or a sub-beat (an eighth note, in most cases), but whether we keep tabs on which subdivisions of said beat or sub-beat the syllables might belong to is another thing entirely. ‘Short’, ‘similar’ and ‘subdivision’ are probably more important experiential categories than ‘dotted eighth’, ‘triplet’ or ‘2:1 ratio’.

In track A6, I discussed the concept of beat bins, wherein inconsistency between placements of rhythmic events expressing the same metrical category in different layers of the composite auditory stream produces a ‘wider beat’. Anything placed within the bin is experienced as part of the beat, regardless of whether we can also ascertain that the events from the different layers are not completely simultaneous. A similar phenomenon arises when metrical anchors and other metrically structuring events from other layers diverge slightly. Due to expected convergence, the specific beat to which the anchor belongs will be widened to encompass the (slightly diverging) events from the different layers. Note that this is slightly different from the analyses of beat bins in the literature. While a perceptual reference structure involving wide beat bins is the result of a consistent and repeated pattern, and the bins stay that way as the pattern continues, the effect here is local. The single beat is widened, but it does not necessarily influence the perceptual ‘bins’ of surrounding beats. It is the result of the structuring power of the flow rather than that of repetition and categorical nuance.
Metrical anchors and rhythmic tolerance: ‘Si Eiga Rås’

In contexts where the musical background is sparse or missing, metrical anchors are given all the power to define the rhythmic categories of the compound musical stream. When a cut removes some or all the elements of the musical background, little remains in terms of rhythmic events to compete with the metrical anchors in dictating beat positions. On such occasions, then, there is the potential for huge durational variation within rhythmic categories. To exemplify this, we will first look at one bar from Norwegian hip-hop group Side Brok’s 2004 track ‘Si Eiga Rås’, illustrated in figure B3.2.

![Figure B3.2: Metrical anchors and rhythmic tolerance in Side Brok’s ‘Si Eiga Rås’, from 1:58.](image)

When the beat is cut after the one, the vocal track delivers the punchline with four clear metrical anchors, indicating the four beats of the musical metre. The durations of each syllable and the groups of syllables, however, correspond poorly to the expected durational categories for which the track has primed us. As shown in the figure, the duration of the individual syllables varies from as short as around 119 milliseconds (significantly less than an even sixteenth note at the track’s 102 bpm pace) to as long as around 268 milliseconds (almost approaching eighth-note territory). The groups of syllables indicated by the metrical anchors (excluding the final, solitary one-syllable beat 4) are also longer in duration than anticipated.

One might expect such significant deviation from quantised durational categories to result in a very audible rhythmic effect—that is, to sound ‘odd’ or ‘different’ in some way. This is not

---

138 Note that these millisecond values are at best approximate, as they are the result of my own plotting of the P-centres of each syllable, using a combination of softwares (Praat and Reaper). The important things is the variation/difference, not the exact values.
the case, however, perhaps because of a ritardando effect, since the beat categories are extended in duration. One would then expect a gradual elongation of the beat and subdivisional categories, though, and that is clearly not the case either. The third beat is shorter than the two preceding it, and there is no clear pattern to the durations of the subdivisions of each beat. Most importantly, I would argue that there is no strong experience of the music slowing down significantly. There is a slight phrase-final elongation, perhaps, but not a major ‘ritardando gesture’.

This lack of clarity in the subdivisional categories is another interesting feature. The groupings of three, four and three syllables, respectively, are indicative of a figure where the first and third groups of syllables are eighth-note triplets and the second group is sixteenth notes, but the complete absence of durational consistency between the syllables means that there are, in fact, no clearly distinguishable subdivision categories in this passage. While one would normally be able to separate eighth-note triplets and sixteenth notes by virtue of their duration alone in a musical context like this (aside from the cuts which happen every eight bars, the musical background presents a clearly quantised musical metrical environment), it only takes a moment to blur these durational categories here. Each syllable belongs to the general category of ‘subdivision’ rather than to the ‘triplet’ or ‘sixteenth note’ categories specifically. The only way to distinguish quadruple from triple subdivisions here involves the metrical anchors. The poetic-rhythmic grouping of stresses (the prosody) takes precedence over any duration-based music-rhythmic grouping. Still, some subdivisional organisation persists. As the eighth-note-triplet and sixteenth-note categories meld together into a single density-referent category, the four syllables and subsequent quadruple subdivision of the second beat mean that there is a weak eighth-note subdivisional category present. The ‘Kå-’ syllable is clearly a stressed syllable even if it is not a metrical anchor. This means that while there is no absolute durational organisation of the subdivisional categories, we have not fully converted to strict prosodic organisation either. Rather than being reduced to a single level of subdivision marked by stressed and unstressed syllables (dupe and triple), the subdivisional categories which meld together are those of triple and quadruple.

139 The ritardando effect is also a trope in hip-hop harking back to the DJ technique wherein one stops the record player’s motor, causing the record to slow down and gradually drop in pitch.

140 This ‘blending’ of quadruple and triple subdivisions is reminiscent of Stover’s concept of ‘beat span’, where rhythmic patterns in Latin-diasporic music can be mapped onto both sixteen-cycle and twelve-cycle metrical frameworks simultaneously, with one or the other dominating at different times or both being equally balanced (Stover, 2009, pp. 148-150).
Interestingly, we are not pulled out of a rhythmic (and metrical) continuum by this sudden variation in beat duration. Instead, it is barely noticeable unless one is actively listening for it—we accept it without question. This phenomenon is accurately described by Mats Johansson as the first element of his concept of *rhythmic tolerance*:

> The flexibility of the rhythmic framework—that is, the fact that bars and beats may vary considerably in terms of both absolute and relative length [duration] from one part of a performance to the next without compromising the experience of flow, tempo and groove. (Johansson, 2010a, p. 69)

Johansson examines this tolerance within the context of Scandinavian folk fiddling, and the variations that he finds in his analysis are, unlike ‘Si Eiga Rås’, both more consistent over longer musical works and features of both beat- and bar-level rhythmic units. The ‘Si Eiga Rås’ example involves a singular occurrence within a larger musical context which is mechanically quantised at the bar level. Still, we have no issue with allowing the (unaccompanied) vocal track and metrical anchors to completely dictate the rhythmic structure at the beat level. Interestingly, it is at the bar level where we experience some categorical confusion due to the durational variation in the rhythmic units. As the three first beats are elongated and the silenced musical background does not adjust accordingly, the final beat of the passage turns out to be significantly closer to half the duration of a full beat in the 102 bpm context than it is to a full beat. Considering that the three preceding beats indicate an even slower tempo context, this means that when the first beat of the chorus hits, it is close enough to the metrical anchor of ‘crack’ that it might be interpreted as being syncopated (placed at the ‘four-and’ position of the preceding bar). Of course, this confusion in the listener, if it occurs at all, is only momentary, since when the chorus unfolds, it is evident that there is no syncopation to start it off. This type of ambiguity—where one probable interpretation is quickly supplanted by another—is a common type of ‘interesting surprise’ in a musical stream. The following example is a more severe type of surprise, wherein the ambiguity lingers and might not wholly resolve even as the stream continues and drags us along with it.

**Metrical anchors creating categorical ambiguity: ‘HiiiPower’**

Kendrick Lamar’s ‘HiiiPower’ (2011) presents a context rich in the heavy analogue modulation of digital rhythmic categories. For one thing, this sample-driven track has elements which are not perfectly quantised to the sixteenth-note-level density referent, but
more importantly, Lamar’s flow is ‘expressively timed’, with a tendency to lag slightly behind the drum track. Towards the end of the first verse, there is a passage—illustrated in figure B3.3—which takes this expressive timing a little further. In combination with the metrical anchors, this creates a very interesting categorical ambiguity.

![Figure B3.3: Metrical anchors creating categorical ambiguity in Kendrick Lamar’s ‘HiiiPower’, verse 1, from 1:23. Metrical anchors marked in green (matching their corresponding expected convergent beats). Syllable P-centres’ temporal positions on a metrical grid corresponding to the musical background marked with red dots. Potential rhythmic figure interpretations in musical notation below.](image)

The passage in figure B3.3 is the concluding statement in a verse preaching self-awareness and reflection with a harsh social commentary, and its final punctuation is the last syllable (the unconventionally pronounced ‘-glyph’s’), falling exactly on the canonical position of the four. The rest of the phrase, however, features the collision of a digital rhythmic feature—an off-beat cross-rhythm—and the analogue feature of ‘behindness’.

As discussed in track A6, one way to quantify the nature of a rap flow’s expressive timing is by measuring its disalignment from some sort of metrical ruler. Using a perfectly quantised ruler (the standard grid of a DAW) at 81 beats per minute, then aligning the backbeat

---

141 Lamar makes it clear that he is not sugar-coating anything by stating ‘you’d die from diabetes if these other niggas wrote it’ earlier in the verse.
142 Since hip-hop tracks tend to be quantised at some structural level, they almost always display a consistent bpm throughout a track. Most of the time, a simple tap tempo measure is enough to find the correct bpm, and in
snare’s transients to this grid, we derive a sensible starting point for looking at the timing in Lamar’s flow. Looking further into Lamar’s flow on ‘HiiiPower’, he clearly trends, albeit to a varying degree, towards lagging behind the beat. This is not done consistently throughout the verse, and the amount of disalignment varies as well, but he rarely ventures ahead of it in any case. This means that one emergent expectation we arrive at while listening to the verse is that variation in the timing of syllable positions will be behind and not ahead. So, when the passage transcribed in figure B3.3 arrives, we are primed to interpret the first metrical anchors (as well as the stressed syllable ‘py-’ in the eighth-note position) as introducing an off-beat phrase starting on the final sixteenth-note position (the topmost notated transcription).

However, there are several factors which oppose this interpretation of the figure. The behindness of the timing means that the off-beat syllables’ positions end up quite close to the canonical on-beat positions. Initially, this is not a problem, but the second metrical anchor—the ‘-mids’ syllable—which is particularly stressed pulls towards the expected convergence of metrical anchor and pulse-beat position. Amongst the reasons for this heaviness is Lamar’s delivery: there is a clear phonological stress, by way of a lowered pitch on this syllable, which further emphasises the already prominent accentuation. Another significant factor is the longer syllable duration. This both allows the syllable to be syncopated in the first place and suggests an agogic accent—one which increases the weight of an already heavy anchor. When this heaviness and the syllable’s disalignment from the grid combine, the syncopation is weakened, making it more likely that one would interpret the syllable as being on the beat.

Thus, it is now possible that some listeners might interpret the segment as being one of the bottom two rhythmic figures in figure B3.3, or they might switch interpretations throughout the unfolding stream from the top figure to one of the others—probably exactly at the arrival of the ‘-mids’ syllable. It is also crucial to keep in mind that our perception process is not as black-and-white as the notated figures, nor does it happen at a perfect real-time resolution. There are questions as to whether we need to assemble coherent figures/reference

the occasional cases where there is a fractional bpm number, there are numerous ways of finding the exact fraction using a DAW.

143 The position of the red dots on the grid in the middle of figure B3.3 is derived from my measured/interpreted P-centres drawn onto the grid of a DAW. A snapshot of this analytical process can be seen in figures like B3.4, where the drum track is used to align the grid with the vocal track waveform and spectrogram.

144 The duration and pitch accent of the metrical anchor are also the reasons why the bottom interpretation (with the metrical anchor being an eighth note) is a valid one. This interpretation is contradicted slightly by the phonological alternating stress pattern, however. The pitch contour shows an alternating high-low pattern on ‘building my own’ which weakens a triple subdivision of ‘building my’. The triple subdivision is still a very intuitive prosodic interpretation, of course, and different listeners will inevitably interpret the structure differently.
structures/schemata in our minds when listening, and if we do, how much mental processing power is allocated to altering such figures when we are presented with new information? A more general sense of tension and release, or ambiguity and clarity, might adequately explain how these passages are experienced, whereas the transcription, representation, analysis and dissection of these minutiae are simply means of exposing different types of rhythmic techniques rather than windows into our perceptual apparatus.

Regarding the specific passage in figure B3.3, there is room for a slight meta-analysis. Having listened to the passage repeatedly, even hundreds of times, both on its own and in the context of its several preceding bars, as well as with the musical background and without it, I must resign myself to the fact that its identity simply changed for me. I cannot hear the passage for the first time ever again. Reflecting on all the ways I have heard it and other people’s impressions of it have left me with some insights, however. (1) It is quite possible to hear both the syncopated and the straight versions, and to switch between the two modes. (2) That said, I keep coming back to the syncopated version as the ‘correct’ one, which stands in contrast with the next point. (3) When evoking an ‘archetypal version’ in my mind, without listening to the track, the metrical anchors overpower the syncopation, and I am left with the straight version. This leads me to believe that it could be fruitful to think about rhythmic figures in the same way we analyse poetic lines. While there might be (and typically is) a clearly dominant perceived figure, there exists ambiguity and alternative figures which are weakly present alongside the dominant one. Hence, it becomes possible to ‘code switch’ between different figures over repeated listenings, and to hear the weaker alternative structures alongside the dominant one as a significant part of the listening experience. One might have to suppress certain parts of the compound auditory stream to justify the weak alternative, but it is nevertheless strong enough to impact the listening experience in a significant way. This ambiguity should be considered aesthetic.

Point 3 has other implications as well. Firstly, the idea of an evoked archetypal figure is at odds with music’s unfolding through time. An archetypal figure implies that the rhythmic passage has been taken out of its temporal context and fixed as a template or mould which can be slowed down, sped up, recontextualised, repeated and reordered. In a way, this is like how one might (but does not have to) interact with the rhythm of printed poetry while reading it, as one is free to stop and consider the words’ rhythmic organisation at any time. With a small epistemological leap, this could suggest an explanation for how prosodic stress might supplant the musical rhythmic and musical metrical information in this archetypal version. Another
way of thinking about it might be that the deadpan (that is, unsyncopated) version with metrical anchors coinciding with the beats of the (now conceptual and fixed in time) pulse is a kind of ‘ultimate digitisation’ of the rhythmic passage. The stressed syllables are the digital categories, whereas the syncopation is a type of analogue modulation of said categories. This is one way in which the concept of cognitive quantisation is useful when one is attempting to fit these musings into the analysis of music as it unfolds in time. In situations where our rhythmic perception is coloured by uncertainty, confusion or ambiguity, we apply a sort of Occam’s razor and reach for the simplest solution—or, perhaps, the first solution which appears. After all, we are working under severe time constraints, and the next task is only fractions of a second away. Whether or not the solution we arrive at is consistent with the preceding totality (or is immediately contradicted by what comes next) is of no particular importance. We apply the explanation which works in the moment and store only the parts we need for later use. In the rhythmic stream of a rap flow, the metrical anchors must be among the most salient types of structuring information.

In summary, then, to explain why the passage in figure B3.3 is so interesting, we must introduce some hypotheticals. Assuming the syncopated figure is the dominant (or ‘correct’) one, a perfectly quantised version of it would be experienced quite differently from the actual version. If there is no possibility of confusion as to whether the disalignment between metrical anchors and beat positions is digital (syncopated/off-beat) or analogue (early on-beat), there is no real ambiguity, and the listener is robbed of the aesthetic experience of their split-second subconscious rhythmic interpretation. There might be no real doubt as to whether the syncopated figure is what we hear, but Lamar’s timing is so expressive, and the ‘-mids’ syllable is so structurally ‘heavy’, that it carries with it much more on-the-beat-ness than a mechanically quantised version would.

Bothness: ‘Eple’

While the passage in figure B3.3 can be thought of as having one clearly dominant rhythmic structure (the syncopated one), and another one which is only weakly present (the deadpan), it is also possible to have rhythmic passages where one rhythmic interpretation is supplanted by another as new information is revealed. Rather than thinking of this as a new ‘correct’ version which invalidates an old obsolete one, it is more accurate to say that both versions are descriptive of what one perceives as the music unfolds. In certain rhythmic passages, there are specific points of interest which can be identified as pivot points between two ‘incompatible’ rhythmic figures—incompatible in the sense that their merging would be impossible to
represent in fully quantised systems such as traditional music notation. Some rhythmic events (syllables, in the context of a rap flow) will inevitably be considered two different categorical rhythmic ‘things’ (rhythmic values, rhythmic positions, parts of a rhythmic pattern, and so on), and this is what I call bothness. A common way for bothness to appear is through timing discrepancies between different layers of a composite auditory stream, as is the case with the example in figure B3.4.

Figure B3.4 Bothness in Neste Planet’s ‘Eple’ (2018), verse 1, bars 5–6, from 0:53. Waveform and spectrogram with superimposed lyrics showing the microrhythmic relationship between defining elements of the musical background (from top: bass drum, rimshot, shaker) and the vocal track (the bottom waveform and spectrogram). Grey vertical columns indicate eighth-note positions. Interpretation of rhythmic structures and a shift between the two transcribed below.

Norwegian rapper Linni uses expressive timing in his flow to create a certain vibe, as in the passage in figure B3.4 from the track ‘Eple’. Talking about his approach to timing in his flows, Linni says:

> It happens naturally, but I know that I do, sort of. But it’s not that I think ‘ooh, now I’m gonna drag it’ when I write. But I just notice how, when I hear it again, I hear like ‘okay, that’s too ABC for my taste’. And since I’ve always thought that... firstly—I have a thing with that when you feel like you’re proficient at something, I believe you should start breaking some rules. If you’re conscious about it. And to me, it also

---

145 An exploration of ‘bothness’ (without the term itself) can be found in Danielsen’s (2015) analysis of ‘Nasty Girl’ by Destiny’s Child (2001), where the rhythmic fabric simultaneously implies both triple and duple subdivisions of the density referent and slow and fast basic pulses, and a synth bass figure presents a cross-rhythm which appears to alternate between triple and duple feels depending on the rhythmic context in which it appears. The affordance of simultaneous triple and duple interpretations is also reminiscent of Stover’s ‘beat spans’ (2009).
contributes to the contents of my songs. I want it to be a little dreamy and ‘swimming’. That all the boundaries should be broken a little. Like, that I both, kinda, play with the genre, the norms in the themes I bring up in my music, and so on, but also, in a way, that I don’t need to . . . I feel like I don’t need to prove that I can rap on the beat, if that’s what it’s gonna be about. I like that it feels like I’m surfing. To me it’s just a way to have a type of control. Which makes it so you can go a little outside the wake when you’re waterskiing, if you know what I mean? (Linni, January 17, 2019)

In ‘Eple’, the musical background is sparse. The drum tracks’ density referent is mostly at the eighth-note level, leaving the vocal track and the occasional drum hits (there is one bass drum and snare hit every two bars which land on what could be considered an off-beat sixteenth-note position) to define the blurry boundaries of a sixteenth-note level. Additionally, there is an incongruency between the bass drum and the shaker in terms of where they position the off-beat eighth notes. The wider grey vertical columns in figure B3.4 indicate a beat bin defined by the distance between the bass drum off-beat and the shaker off-beat. In fact, the different drum tracks are programmed using different swing ratios, muddying the compound subdivision categories still further: the shaker has a 9:7 ratio for eighth notes, while the syncopated rimshots (where there is never a successive on- and off-beat hit, effectively hiding the categorical swing pattern by spreading the events out) are programmed to a 5:3 ratio at the sixteenth-note level. This rhythmic environment allows for a great deal of freedom for Linni’s vocals to be ‘dreamy and swimming’ in the placement of syllables, and the passage derives its ambiguity and bothness from a mismatch between the internal relationship of the syllables and their positioning relative to the musical background.

The first part of the phrase ‘ta det dag for dag, ting’ [take it day by day, thing] has a stress pattern of strong-weak-strong-weak-strong plus a slightly ambiguous final syllable which could be considered stressed if not for the alternating pattern priming us for it to be weak. The most intuitive interpretation of this part of the phrase is that the first syllable is a late (analogue) on-beat (digital) rhythmic event. Its position relative to the drums is late enough that it could conceivably be interpreted as being displaced one position on the sixteenth-note level, but this would result in a consistent mismatch between stressed syllables and weak

---

146 I have attempted to translate as faithfully as possible from the original recording. However, some Norwegian expressions and words lack a direct English equivalent, and in those cases, I have tried to preserve meaning rather than supply the closest word-for-word translation.

147 Producer Erlend ‘Kvam’ Lyngstad works with FL Studio and tweaks the microtiming profiles in individual tracks using the DAW’s swing ratio parameter (Kvam, interviewed by the author, September 28, 2018).
sixteenth note-positions, and the passage does not feel like it features off-beat phrasing. At the same time, the syllables are roughly similar in duration, leading us to experience them as being placed in consecutive sixteenth-note positions, as in the topmost transcription in figure B3.4.

Complicating matters further are the multiple clues that the sixth word (‘ting’) is in the ‘two-and’ position. For one thing, both its absolute temporal position and its alignment with the drum track are quite evident from the figure (and present in the listening experience). In addition, as previously stated, the only valid reason for it not to be stressed is the alternating pattern of stresses which precedes it. These factors advocate for an emergent rhythmic interpretation for this passage—one which will be promptly reinterpreted when we are faced with contradictory evidence. There is every reason to hear the opening of the phrase as a late on-beat arrival, but when the first ‘ting’ appears, its identity as a ‘two-and’ is what anchors our listening experience. I do not believe that this forces us to reinterpret the entire passage, however. It is clearly not an off-beat structure, and the distance between (and vocal delivery of) the words ‘dag, ting’ does not justify an interpretation of them as even eighth notes (as notated in the bottom transcription in figure B3.4). Instead, our localised reinterpretation simply implies that the key syllable—‘ting’—is both things. The figure morphs from the one to the other without the need to reconfigure the entirety to make sense. Note also that Linni does not rein in the dreaminess of the passage after ‘dag, ting’. ‘Så ting’ and ‘vanning’ might be more off-beat-on-beat (the bottom interpretation) than on-beat-on-beat (the top variant), but the off-beat syllables are quite early, as can be seen by how much earlier ‘så’ is than the rimshot track’s marking of the sixteenth-note position in the figure. Ultimately, the passage reveals an evolution from one dominant rhythmic structural interpretation to another, with the single pivoting rhythmic event being one thing in the initial interpretation and another in the emergent reinterpretation. This event’s bothness is central to the passage’s entire character.

**Off and on-beat and stretched subdivisions: Bothnesses in ‘Do You Remember’**

Another example of syllables displaying contradictory categorical information is a passage from Chance the Rapper’s ‘Do You Remember’ (2019), visualised in figure B3.5.
Again, contextual clues from the ongoing musical surroundings prime us for certain rhythmic interpretations. As can be seen in the figure, the instrumental tracks leave the first beat open for Chance’s flow to dictate the initial rhythmic categories. Additionally, every line starts either with a pickup (bars 1 and 4) or on an off-beat after the initial downbeat on the one (bars 2–3, 5–6). Thus, we experience a bothness in the opening of the line which is quite evocative of but opposite to our previous example. Rather than first hearing ‘used to rock’ as a late on-beat structure, we are primed to interpret it according to the rules laid out by the entire track to this point as starting after the downbeat as an off-beat structure. Since the first syllable falls well after the (silent) one position, this interpretation is initially validated. However, it takes less than a second for the metrical anchor of the first ‘South’ to cast it into question. The word ‘rock’ is too long in duration to be a sixteenth note, and ‘South’ is a heavy on-beat syllable (a digital quality), however late (an analogue quality) it might be. Yet, our experience of the beginning of the phrase is still that it starts after the one. Thus, the bothness experienced here is not that one syllable belongs to two different categories in the two morphing interpretations but rather that a group of syllables (‘Used to rock’) does. They are initially off-beat, stemming from their temporal position and the contextual reference structures offered by the preceding music, and then they become on-beat, as the metrical anchors retroactively identify them as a figure of two sixteenth notes and an eighth note spanning the full first beat.

---

148 This is the only place in the song where the bouncy, syncopated Rhodes, the main accompanying instrument, is muted completely at the beginning of a bar. My guess is that it is a production choice intended to avoid clashes with the expressive timing of Chance’s flow in this passage.

149 Both instances of ‘South’ are capitalised because the first (probably) refers to the clothing brand ‘South Pole’, and the second refers to the community area in Chicago known as the ‘South Shore’.
The alliterated metrical anchors of ‘South’, ‘Summers’, ‘six’ and ‘South’, respectively, are strong markers of the experienced beat positions, and the way in which they are inconsistently timed sets the stage for yet another instance of bothness in the figure spanning the third beat of this passage. Three out of the four metrical anchors are late, the first less so than the last two, which are late by a durational equivalent of a sixteenth note.\(^{150}\) The third metrical anchor, however, is precisely on the grid indicated by the instrumental track. This could indicate two different things: either Chance uses an inconsistent expressive timing (in the sense that it does not display a consistent relationship to a grid, not that it is unplanned or a mistake of any kind) or he is consistently approximately a durational equivalent of a sixteenth note behind the instrumental track. The latter case would explain both the opening syllables being ‘on’ despite being late and the upcoming bothness, but I would argue that the first explanation—Chance using expressive timing which is not consistent in its directional pull on the rhythmic categories—is clearly the one which more closely matches our listening experience.

The part of the phrase spanning the third beat (‘summers taking the’) consists of five syllables of similar duration, and this is what creates issues for any commonplace explanation of the phrase’s rhythmic structure. If it were merely a case of five equal rhythmic units subdividing a larger beat-level unit which was uniform with its surrounding beat-level units, the rhythmic interpretation would be relatively simple—that is, it would be a quintuplet. However, as should be clear from figure B3.5, the span these syllables fill is not of a uniform beat-level unit, nor do the syllables have durations which imply a beat subdivision into fifths. Instead, the syllables appear to be indistinguishable as durational units from sixteenth notes (as the latter have been established throughout the verse). The upshot of this is that the beat span is filled by five equivalent rhythmic units which are sorely lacking in ‘quintupletness’ and must also be considered sixteenth notes—which is an impossibility. The prosody presents another interpretation which is more intuitive than the quintuplet one—a 2+3 grouping of the syllables indicating two sixteenths and a triplet. However, this is also undermined by the strong sense of all five syllables being ‘sixteenth note–like’ in duration.

This ‘impossible’ outcome is what lends credence to the interpretation that the flow displays a consistent lateness in timing rather than a more generally ‘expressive’ timing. If this were the case, the rhythm would not be impossible to represent within a quantised system—it would

\(^{150}\) Note that I refer to it as the ‘durational equivalent of a sixteenth note’ rather than ‘a sixteenth note’ to emphasise the difference between the durational span and the categorical unit.
simply span an extra sixteenth note, with the first syllable (‘sum-’) placed on the last sixteenth note position of beat 2 rather than the downbeat position of beat 3. However, this is not a satisfactory explanation either. There is very little in the music which supports an off-beat interpretation other than the fact that it makes the math of the transcription add up. The two stressed syllables of the five-syllable group (‘sum-’ and ‘tak-’) simply do not feel off-beat, and the sheer weight of the metrical anchor (which is strengthened by the alliteration pattern of the surrounding anchors on the downbeat positions) pulls the listener away from any other interpretation than ‘sum-’ landing on beat 3.

Again, I turn to bothness as an explanation—this time, a bothness of subdivision. The rhythmic structure as indicated by the prosody (two sixteenths and a triplet) is probably the most intuitive interpretation of subdivisional structure, and simultaneously each event is experienced as sixteenth notes in subdivisional duration. The last three syllables are both a triplet and straight sixteenths at the same time. They fit into the rhythmic metrical framework (both prosodic and musical) by their tripletness, and they fit into the even rhythmic unfolding flow as detached from metrical rules in their sixteenthness.

**Do we really hear this stuff? Epistemological considerations**

This track has explored concepts that range from the fairly obvious—metrical anchors—to the more esoteric—bothness. While the former might be an intuitive explanation for how an expressively timed vocal stream might dictate the compound metrical position of an auditory stream, the latter is a complicated description of some types of momentary rhythmic confusion for a listener as composite auditory streams unfold with various layers contradicting one another—that is, with some events displaying multiple ‘incompatible’ categorical belongings. We hear metrical anchors, but we do not necessarily experience bothness in the manner it is described in this track. Among other things, the duration of its presence is too short. For a thesis so focused on what happens in time rather than on out-of-time mental representations, it might seem paradoxical to outline a long-winded theory of a phenomenon which in practice is but a snapshot of an ephemeral moment.

There are some epistemological considerations to mention here. Yes, the actual experiential dimension of various categorical ambiguities at the microrhythmic level can probably be more than adequately described as a progression from clarity to confusion to clarity—or, to make the inevitable parallel to functional harmony, from tension to release. Still, because I set out to describe techniques which rappers employ in their flows, there is a need to theorise and
distinguish different facets of this ‘tension’. It is not necessarily about the listening experience anymore but about the difference between the various ways a rapper can use expressive timing to create ambiguities, confusion, tension and release by way of the expected convergence between layers of the composite auditory stream, digital rhythmic categories and the analogue modulation of said categories. It becomes necessary, in the end, to discuss the aspect of production in addition to the aspect of reception.

First, let me make it clear that I believe that the most common origin of these types of expressive timing-related effects is experimentation, coincidence or, at the very least, vague notions rather than specific plans. A rapper might focus on ‘attitude’ or aspects of delivery rather than perfect synchronisation with the musical background—or, as Linni said, they might just see where they end up by water-skiing outside of the wake. One navigates by feel while recording—trusting that one knows when it is just right. A part of the ‘production’ of a rap flow is also its post-production, where one (particularly in the age of computer-based production techniques) can alter the timing of a vocal track in significant ways. Producer trio Basmo Fam explains:

[Some rappers’ flows] are just a little ‘unrhythmical’, which can be dope. It just gives a kind of attitude, or ( . . . ) an emotion that can be cool. Even if it is ( . . . ) a little unmusical in a way. ( . . . ) It is rare that people try to be [unrhythmical], it is more by accident, but many choose to keep [the take] that has the most attitude ( . . . ) independent of [the timing]. ( . . . ) But it happens that we tighten it up, patching it up after. ( . . . ) Sometimes you choose to keep it, but most of the time you kinda want to make it a little closer [to quantisation], because if it is some small things that deviate, then that’s fine, but if it’s, like, offbeat all of the time, it gets tiresome to listen to.

(Basmo Fam, September 17, 2018)151

Rapper/producer Oral Bee describes a practical application of altered timing in post-production when discussing his recording process:

I refer to myself as ‘one-take-Beezy’, but often there might be a couple of warm-ups [takes], and then, particularly when I work with Chris [Chris Lie],152 as he’s quite anal about these things, we might do three-four-five takes of the verse. And then, none of

---

151 The parts of the interview quoted here are cut together from a passage where all three members chime in on a discussion of rappers’ timing in the studio.

152 Kristoffer ‘Chris Lie’ Lie collaborates with Oral Bee in the duo Splash Bros. and is the ‘house mixing engineer’ at Oslo Records.
them are necessarily bad, but we go in and find the best parts from the different ones. (. . .) It can be single words, two lines (. . .) It could be the melody (. . .), it could be the way the voice sounds, whether it breaks up, or if there are any super-sharp s’s (. . .). And the most important is maybe the attitude of the take, that you sound like a whole together with the beat, that it doesn’t stick out, unless that’s the point. That you kinda sit in the beat. (. . .) And that is a timing thing too. It happens that ‘okay, here I was either too fast or a little too slow with the ending’, and then we go in and take the final word and just nudge it. Then we get it where we want it. (. . .) It might happen that Chris looks a little at the grid [in the DAW] (. . .), but first and foremost it is a feel thing. (Oral Bee, January 25, 2019)

These reflections indicate that in production there is an active curation of which instances of expressive timing are kept and which are either removed or doctored—the post-production stage, that is, represents a central part of the creation of a flow’s microrhythmic signature.

Another implication of the use of ‘comping’ (the process of cutting together parts of different takes) is that some types of quite advanced timing techniques might be easier to accomplish than if a rapper had to rely purely on a single ‘live’ performance. It can also be a boon to the creative process, where one could feasibly compile timing strategies from different passages of a flow from different takes and settle upon a combination of timing techniques which one might not have come up with outside the specific context of digital recording and editing. A speculative example which comes to mind is Kendrick Lamar’s first verse on The Game’s ‘On Me’ (2015, from 0:43), analysed in Komaniecki (2019, pp. 57-59), where Lamar alternates between straight and swung sixteenth notes, morphing from the former to the latter. Without knowing how the recording sessions unfolded, there are several creative scenarios which seem plausible here. Perhaps Lamar did different takes, some with straight and some with swung sixteenth notes, and the final version was comped from among them? Perhaps the idea of switching between straight and swung came up from listening to these different takes? Or perhaps the rappers came to the session with ideas for their verses already worked out, or even with recorded demos, only to discover that they had different approaches to the sixteenth note swing (The Game employs swung sixteenth notes)? Of course, this is only speculation, and Lamar could well have planned or improvised the passage without any of these factors influencing him. Nevertheless, the possibilities that the modern working environment offers for audio manipulation enable the artist (or the audio engineer or the former through the latter) to experiment with timing in ways which are practically unfeasible. It would be horribly
impractical to do multiple full takes of a full phrase where everything is exactly the same save for the temporal position of a single syllable, but in any contemporary DAW one can create, listen to and compare such versions in just a few seconds. The tools available influence the end product in significant ways, and a rapper’s flow is therefore the product of not only a physical performance captured by a microphone and recording equipment but also the post-production process and its possible comping of various takes, stacking of dub tracks and so forth.

The communication and representation of these types of microrhythmic features in various contexts add up to another epistemological and ethical can of worms. I will return to this in the final track of the thesis (B5). For now, I will shift my focus to another type of conceptual ‘bothness’—rap’s double relationship with song and speech.
Track B4: Speech-song or song-speech?

As might be deduced from its slightly tautological title, this track dives into rap as a type of **vocal style**, allowing the constituent of **delivery** to take the front seat from, or at least ride shotgun alongside, the flow. I will explore rap’s relationship to the two poles (or perhaps ‘siblings’) of speech and song, engaging in turn with both how this has been and how it can be analysed. While this means that there will be analysis of both rhythmic and melodic content, the latter will primarily be viewed in terms of its impact on rhythm and form. This raises questions about **what rap is**, and about the relationship between delivery and flow. However, a core point of this track—one which will be made clear in the longer analyses towards the end—is that the landscape between song and speech is resistant to cartography. Taking advantage of this fact, rappers utilise the quality of being in-between these poles to enhance their flows in various ways, some of which will be exemplified in this track.

**Speech-song: Rhythmicity**

Both speech and song are featured as references in Krims’s taxonomy of flow styles (discussed in track A1). He describes both the ‘sung’ style as well as the ‘speech-effusive’ style (as one of two ‘effusive’ styles, alongside ‘percussion-effusive’). Interestingly, throughout his descriptions and comparisons of the styles, Krims’s focus remains on various rhythmic parameters. Sung style is characterised by what this thesis dubs convergent metrical structure\(^\text{153}\), repetition of rhythmic figures and on-beat accents, whereas Krims sees the effusive styles as more complex, using (what I would describe as) divergent metrical structure and creating crossrhythms and counterrhythms through the off-beat placement of syllables. One question that comes to mind, then, is whether these would be identifiable differences between rap flows which are ‘more song-like’ and ‘more speech-like’. Krims distinguishes between ‘speech-effusive’ and ‘percussion-effusive’ mainly via articulation and features of delivery but also certain rhythmic parameters. A percussion-effusive style displays ‘pointed articulation’ and ‘punctuate[s] the musical texture’, whereas a speech-effusive style disregards this rhythmic clarity ‘with little sense often projected of any underlying metrical pulse’ (Krims, 2000, p. 51). Perhaps, then, while convergent metrical structure would be a sign of song-likeness, divergent metrical structure is not speech-like in and of itself but instead less song-like?

\(^{153}\) In Krims’s words: ‘strict couplet groupings’ (Krims, 2000, p. 50).
One can perhaps modify the terminology slightly from ‘speech-like’ and ‘song-like’ to ‘speech-rhythmic’ and ‘song-rhythmic’ to stress that it is the rhythmic features of these approaches to flow which are being described. ‘Speech-rhythmic’ as a term is central to Ohriner’s article ‘Lyric, rhythm, and non-alignment in the second verse of Kendrick Lamar’s “Momma”’ (2019c), and he contrasts it with ‘music-rhythmic’ (as opposed to my Krims-derived suggestion of ‘song-rhythmic’). Whereas Krims’s differentiation between speech- and music- (song-) rhythmicity also encompasses features of the macrorhythmic and quantised zoom levels, Ohriner measures the degree of speech-rhythmicity by the amount of disalignment from a quantised sixteenth-note grid, the argument being that music-rhythmicity is characterised by (relatively) equal rhythmic categories of beats and subdivisions, whereas speech does not exhibit this regularity. Ohriner discusses how disalignment between syllables and quantised surroundings can be understood as ‘out of phase’ (consistently behind or ahead—that is, a systemmatic variation, as discussed in track A6), and also how specific timing profiles are typical of a speech-like pattern. The prime example of the latter is ‘continuous deceleration’, though one might wonder whether disalignment due to continuous deceleration also could be explained as music-rhythmic ritardando? Several of the examples from the previous track should make it clear that not all microtiming disalignments are necessarily markers of ‘speech-rhythmicity’. Or, put differently: even though speech is characterised by not having perfectly regular syllable durations, not all deviations from regularity are speech-like. Ohriner touches on this point in his chapter on expressive timing while acknowledging the more phenomenological nature of the speech-likeness in Lamar’s flow: ‘Some of these lines, to my ears, sound more like the rhythm of speech than that of music’ (2018, p. 17). Ohriner’s observation is relevant, but even though Lamar’s flow is reminiscent of speech’s rhythm, this does not necessarily mean that one can measure speech-rhythmicity in general through the amount and degree of disalignment. There is also the potential for an unsystematic disalignment in microrhythm to not be speech-like. If, as I would argue, microrhythmic features do not enable us to confidently separate speech-likeness from song-likeness by themselves, we are dependent on triangulating these features with something else. And while I have proclaimed and defended at length the fact that it is not necessary to do so when exploring the rhythms of rap, I will now take the shallowest of dives into an enormous topic—that of pitch and melody in rap—as it is key to unpacking the concepts of speech-likeness and song- (or music-)likeness.
99 problems, and is pitch one?

In addition to discussing speech-rhythmicity, Ohriner has also explored the topic of pitch in rap. In ‘Analysing the pitch content of the rapping voice’ (Ohriner, 2019a), he uses tools typically employed in the field of linguistics\(^{154}\) to analyse tracks by various rappers and compare their pitch content with both one another and with archetypical speech and song. A central point is that rap is different from both speech and song, and he supports this first through the citation of various scholars and performers, then by intonational analysis. Ohriner identifies several features where rap is different from speech and explores rap which is song-like or (exemplified using Canadian rapper Drake) fully sung. In fact, one might well interpret some of Ohriner’s findings as an indication that the difference between rap and song is much less clear-cut than the difference between rap and speech.

Another, purely descriptive approach to analysing pitch in rap is that of Komaniecki. Rather than measuring the pitch content of rappers’ voices, Komaniecki limits his analyses to categorical descriptions and transcription of discrete musical pitches. He proposes a spectrum of vocal pitch techniques ranging from ‘more speech-like’ to ‘more song-like’ wherein the former display ‘increased pitch instability’ and the latter display ‘increased pitch stability’. Just as Ohriner’s scale of speech-to-music-rhythmicity situates rap as an in-between form, wherein a track/performance can be more or less speech-rhythmic, Komaniecki’s spectrum puts rap in the same in-between space in its application of intonational techniques. While both scholars acknowledge and illuminate through analysis the fact that rappers can (and do) vary immensely in where on these spectra they position themselves—even within a single track or verse—there abides the clear implication of a categorical difference between the poles of the scales—‘sung’ and ‘spoken’—and the in-between-form of ‘rapped’. Ohriner, in his analyses of pitch, also present features which function as evidence for this difference, particularly in the relation of rap (and song) to speech. However, I believe it is crucial to stress that not all rap displays this in-between-ness, and that not all in-between-ness is the same. While rappers do manipulate categorical ambiguities between song and speech and rap, the categories are not necessarily as easily separable as one might think.

One seemingly paradoxical argument for my stance can be found in Norwegian linguist Jan Hognestad’s analysis of rap vocal tracks (Hognestad, forthcoming 2022). When describing the

\(^{154}\) Such as the softwares ‘Praat’ and ‘Prosogram’, the former of which I have occasionally employed in this thesis as well.
small collection of rap tracks on which he performed his preliminary analysis, he makes the following observation:

[Most of the available tracks] were evaluated as less interesting than these two [tracks analysed in the article], simply because they consistently consisted of undisputable song. (. . .) To the degree that this might be representative, it could imply that rap is actually sung to a larger degree than the international academic literature would indicate. (Hognestad, forthcoming 2022, my translation)

If the tracks are both ‘undisputably sung’ and clearly considered ‘rap tracks’, does this not imply that ‘rap’ is something other than simply a vocal discipline somewhere on a spectrum between speech and song? Can something both be fully ‘sung’ and ‘rap’ at the same time? I believe the answer to this question is a resounding ‘probably’. There are not many categorical intonational differences between song and rap identified by Ohriner, Hognestad or Komaniecki; instead, there are degrees of ‘more song’ and ‘less song’. The use of few(er) discrete musical pitches seems to be a feature of rap, but pop tunes with a severely limited melodic range are by no means uncommon. Is it at all possible, then, to truly separate rap from song by analysing intonation? Hognestad makes an interesting point when discussing the differences between speech and song:

On the one hand, we are dealing with a dichotomy where the differences are of a rhythmic and melodic character. On the other, there is an informal tendency to overrate the distance between the forms of expression of speech and song, since [both] are fundamentally characterised by rhythm and melody. (Hognestad forthcoming 2022, my translation)

This relative similarity of the two opposite poles between which rap is placed seems to invite the idea that the categorical boundaries are, at best, blurry and malleable. Additionally, both Hognestad’s and Ohriner’s analyses end in discussions which serve to challenge the generalisability of their findings. There are nuances and degrees to the intonational features which define rap, and one might question whether it is sensible to even attempt to generalise about this at all, or whether ‘rap’ as a vocal practice can be adequately defined through its measurable constituents like pitch and rhythm.
Applying the ear test—is it rhythm, is it intonation, or is it Something Else?

Hypothesising that the categorical difference between speech, rap and song is something that is more phenomenological than directly measurable, I will again turn to a search for negative evidence. While we are clearly able to listen to a short excerpt of music and assess whether it is speech, song or rap, it is not necessarily possible to measure what makes these categories identifiable. Even features identified as ‘evidence’ for categorical difference might not be sufficient in and of themselves.

Examples of rap being nearly indistinguishable from speech can be found in British rapper Mike Skinner’s verses on the first Streets album, Original Pirate Material, from 2002. Particularly on the track ‘The Irony of It All’, where Skinner plays two opposing characters championing alcohol and marijuana, respectively, he employs both intonational and rhythmic features which are evidence of speech rather than rap, going by Ohriner’s categorisations, including upward glissandi (intonation) and deceleration towards line endings (rhythm). One might well characterise Skinner’s vocal performance on ‘The Irony of It All’ as speech, not rap, but the rhythmic organisation is much more regular than one would expect of speech and would seem to indicate what Ohriner would call ‘speech-rhythmic rap’, which is still ‘rap’. In another track on the same album, ‘Don’t Mug Yourself’, Skinner employs clear speech-intonational features but to a lesser degree. The syllables are much more evenly spaced in ‘Don’t Mug Yourself’ than in ‘The Irony of It All’, and the intonation is closer to what Hognestad would describe as ‘rap intonation’: rather than displaying two relative intonation levels of high (H) and low (L), the flow revolves around a third middle (M-)tone. This M-tone, which Hognestad refers to as a ‘recitation-tone’\(^{155}\) (Hognestad forthcoming 2022), is more stable in pitch than what is common in speech. Skinner even occasionally imitates melodic content from the musical background rather than veering into more speech-like intonation. An exception which makes the difference in intonational approach clear appears in the passage preceding the second verse (0:49–0:56), where there is a spoken interlude of sorts. When the verse begins, in turn, there is a stark difference, both rhythmically and melodically, from the spoken passage preceding it.

One might wonder whether ‘fully spoken rap’ is even a theoretical possibility. The difference between prose and poetry—the division into lines—also indicates a difference in how we

\(^{155}\) A nifty point that is lost in translation (or perhaps rather made possible by the codification of the untranslated Anglicism ‘flow’ in Norwegian rap discourse) is that Hognestad describes this M-recitation tone as ‘flowens flytnivå’, which directly translates as ‘the flow’s flow level’.
organise segments of text. While prose and speech are segmented through syntax and intonation into sentences and phrases, poetry has lines (which, as thoroughly discussed in track A5, might correspond with sentences, phrases or the like, or might not). Metrical and/or rhymed poetry, like rap, has other rhythmic and phonological organisational features which impart some song-like qualities to the linguistic structure. This could indicate that one difference between speech and rap involves the way in which the segmentation is organised. Or can we simply say that if there is no rhyme, it cannot be rap? While I am not aware of any well-known instance of non-rhyming rap (and the idea seems alien to me), approaching the idea from the opposite direction and evaluating something that rhymes but is (arguably) not rap could be a way to ascertain whether there is a clear categorical boundary which hinges on rhyme. Norwegian electro-pop, hip-hop-adjacent artist Emile The Duke performs one step further from rap and towards speech than the aforementioned Mike Skinner. Listening to his 2015 track ‘Fredag’, and especially the section (I hesitate to call it a ‘verse’) from 1:07 to 1:38, I encounter a performance that is spoken, not rapped. However, starting at 1:20, he does establish a primary rhyme complex, suggesting a possible poetic lineation and pulling the delivery further away from regular speech. Yet, the section remains spoken in character, possibly because the rhymes are not placed regularly so as to create any sort of more ‘musical’ regular structure to compete with the linguistic structure. Additionally, the rhymes, both in the way they are performed and in their monosyllabic nature, do not rhyme rhythmically. Perhaps rhythmic rhymes are then a necessity for rhymed speech to be rap?

Alternatively, is it just as much a question of contextual clues? The musical background, the delivery, the lyrical content and the artist’s look/image/other tracks are perhaps what makes one consider ‘Fredag’ spoken and Skinner’s ‘The Irony of It All’ rapped (if one even considers the latter rapped). If indeed ‘rap’ is reliant on a collection of features with very few (if any) obligatory ones (that is, more of a gestalt than a physically measurable category of vocal performance), one would expect that extramusical information could be part of what makes something ‘rap’ or not. Pivoting to the other boundary again—the one between rap and song—the case of Bone-Thugs-N-Harmony is another interesting listening test. Their perhaps biggest hit, ‘Tha Crossroads’ (1995), which netted them a Grammy award in 1997, displays the types of vocal performance Hognestad classifies as ‘undeniably sung’, yet the group is considered a gangsta rap group, not ‘r&b’ (the typical classification for music with sung

156 Whether one considers it one long complex with a consonant shift from /k/ to /t/, or two different complexes is of no real consequence to the analysis.
vocals and stylistic elements from hip-hop).\textsuperscript{157} In addition, that Grammy award was in the category ‘best rap performance by a duo or group’. The verses in ‘Tha Crossroads’ are not merely centred around discrete musical pitches but also display a significant melodic range. Additionally, there are several instances of melismata (exemplified in figure B4.1) wherein a single syllable encompasses several notes, something that is a virtual impossibility in an archetypically ‘rapped’ verse.

Figure B4.1: Melismata (red highlight) in Bizzy Bone’s verse on Bone Thugs-N-Harmony’s ‘Tha Crossroads’ (1995, from 0:25).

Is it viable, then, to define ‘Tha Crossroads’ as sung, not rapped? What would be the argument for not doing so? There are some flow-technical points of disputation. The rhyme density is relatively high and clearly higher than in most pop songs.\textsuperscript{158} The rhythmic density itself is also high, and there are more syllables per beat than one would expect from a non-rap-song. There are crossrhythms and syncopations galore, as well as intricate variations in subdivisions with switches between duple and triple. There is also ambiguous lineation, featuring syntax, rhyme positions and rhythmic parallelisms which diverge from one another.\textsuperscript{159} The question, then, is whether all these rap-typical features are enough to define the verses as ‘rapped’. If that is the case, ‘rapped’ clearly cannot mean ‘not sung’, as it is ludicrous to say that the verses on ‘Tha Crossroads’ are not sung. In turn, the evidence here appears to indicate that either ‘rap’ and ‘song’ are not mutually exclusive or one can have ‘rap verses’ with a ‘sung delivery’. If there is any real difference between the two alternatives, it is

\textsuperscript{157} Not to mention the artist’s skin colour, which is sadly, but undeniably, a huge factor in the media’s (and other institutions like the Recording Academy’s) classifications of musical genres.

\textsuperscript{158} In figure B4.1, the primary rhyme class of /eɪ/ is scattered throughout the lines. The ‘end rhymes’ on the fours are emphasised melodically and rhythmically, but there are also rhyme instances on the half lines (‘AIDS’ and ‘played’) as well as with an added one-rhyme (‘prayed’). Bizzy Bone also snuck in the rhyming diphthong in his pronunciation of the final syllable of ‘destiny’.

\textsuperscript{159} An interesting side point here is that perhaps the addition of melodic parallelisms is a factor in determining the metrical structure. When both rhyme position and melodic structure indicate the same lineation alternative, syntax seems to be weakened as lineation evidence.
probably more discursive than practical—another argument for the importance of extramusical features in the definition of ‘rap’. As long as it is difficult to categorically separate (some) rap from song, I believe the inevitable conclusion is that ‘rap’ is more than just a vocal discipline, since evidently there are examples where there are no meaningful measurable differences between rap and song. The difference between them resides instead in factors outside of the vocal track/performance, such as the musical style or genre or the image and narratives surrounding the artist.\textsuperscript{160} However, these song-like examples of rap are edge cases (albeit epistemologically interesting ones), and it is often much more evident—in a measurable sense—that rappers manoeuvre towards speech-likeness or ‘less song-likeness’ in their vocal performances.

**Utilising the in-between 1: ‘Casket Pretty’**

The observation leading to this convoluted discussion about rap in relation to speech and song is that rappers utilise features of this in-between-space to create rhythmic and formal effects in their flows. However, the techniques employed can be very different from one another, and they can be employed in various ways. The rest of this track will be devoted to analysing two such approaches, beginning with Noname’s ‘Casket Pretty’ (2016).

A quite short track with a backing track that employs very sparse drums and a metrically irregular electric piano chord progression, ‘Casket Pretty’ paints a dark picture of life in Chicago, including narratives of death, police violence and the difficulty of coping. The musical metrical structure is probably best described as a repeating unit of four bars of 3/4 plus one of 4/4, together comprising a section spanning 16 beats, or the same length as a regular four bars of 4/4. Noname’s vocal delivery follows upon the rhythmic ambiguities and irregularities in the musical background and employs a variety of *speech-rhythmic* techniques, but her speech-rhythmicity is not necessarily of the microtiming-disalignment variety described by Ohriner. There are certainly both accelerandi and decelerations of the relatively regular triplet stream,\textsuperscript{161} as well as passages which deviate from temporal quantisation in other ways (syllables also generally tend to lag behind the beat slightly). On the other hand, there are also passages where a microrhythmic zoom reveals that Noname is more regular than her performance feels and sounds, as is evident in figure B4.2.

\textsuperscript{160} For example, Bone-Thugs-N-Harmony released their 90s albums on the record label Ruthless Records co-founded by gangsta rapper pioneer Eazy-E of NWA fame.

\textsuperscript{161} The track could easily be interpreted as being in 9/8+12/8 as well. However, I choose to abide by the common rap denominator of quarter notes in this analysis.
As can be seen in the figure, Noname places the syllables of the metrical anchors (‘searching for God in the bottle he gave me’) in the first bar of the verse relatively close to the quantised beat positions, as indicated by the backing track and the graphic grid. While the (triple) subdivisions are slightly less regular, there is no real rhythmic friction between musical background and vocal track to emphasise this condition. Since we do not have any quantised grid with which to perceptually compare the alignment, it is unlikely that we would perceptually experience a speech-rhythmic microtiming disalignment.

---

162 Note that the tracks are not perfectly separated, and there are remnants of the backing track in the vocal tracks. This is because the track was split using the ‘Spleeter’ source separation algorithm from https://github.com/deezer/spleeter, which is fantastic but not perfect.

163 The graphic grid in figure B4.2 is set to sixteenth notes, but there is no analytical point to be made by adjusting the grid to triplets anyway—as we are not capable of hearing a tiny disalignment from these small subdivisional categories without any reference. It is again a question of which ruler one applies, and I am interested in the perceptual, not the mechanical one.
There are other speech-rhythmic features at play at both the macrorhythmic and quantised zoom levels, however. Compared to song or metrical poetry, the phrasing of speech tends to be very irregular, and this type of speech-rhythmicity is clear from the various types of lineation evidence in ‘Casket Pretty’, as can be seen in figure B4.3.

Figure B4.3: Various structuring parameters in the first part of the verse of Noname’s ‘Casket Pretty’ (0:33–1:06).

There is enormous variation in the relative lengths of the segments of text in this excerpt. Looking at the red caesurae which mark Noname’s pauses/breaths, there are some segments which are short phrases—like ‘And we watch the news’ and ‘Too many babies in suits’—and some which go on longer the expected, spilling over both syntactic boundaries and rhyme complexes—like the phrase beginning with ‘Ashes to ashes’ and ending with ‘afraid of the dark’. The latter even complicates its ending, since the next short phrase is clearly a continuation of the poetic content, where ‘dark’ and ‘blue and the white’ are hues to fear. The phrasing is inconsistent in the way it lines up with the musical metre as well. Sometimes pauses and barlines converge (red caesurae and green lines in the figure), but they rarely do so on the boundaries of the metrical blocks (the solid green lines). Rhyme complexes are interrupted too, as when the /-aɪt/ complex appears to give way to a new primary rhyme in ‘goodbye’/‘outside’, then returns again (‘right’ at the end of the penultimate line in figure B4.3). But it is perhaps more of a rhyme shift than a change in rhyme class, as the two rhyme classes are connected through final syllable assonance (the diphthong /aɪ/), and the two-syllable rhyme of ‘goodbye’/‘outside’ might be experienced as rhyming weakly with ‘tonight’. To preserve this ambiguity, all rhymes are marked using the same coloured outline in figure B4.3. Repeated identical or near-identical words are marked as different from

---

164 Interestingly, these two short phrases, though rather distant from one another, might well be experienced as rhyming. In a poetic analysis on the page, the parallelism is clear—both phonologically (with the assonance rhyme between ‘news’ and ‘suits’) and in that the phrases are short and spaced relatively evenly within the larger poetic form. It would appear that the verse was written as a poem before being adapted into/onto a flow.
rhymes in the figure as well, even though some of them rhyme with other words outside of their immediate proximity. It is not possible to authoritatively distinguish between ‘rhyme’ and ‘local repetition’ here. Lines (if one can even call them that, given all their ambiguity) ending without rhyming also complicate the form—the first one (ending with ‘gave me’) and the long one ending with ‘afraid of the dark’, for example, both have rhyme instances within the line but not at the end. The line ‘Where’s love when you need it’ is even devoid of rhyme connections with anything either preceding or following it.

The irregularity and asymmetry of some sections are other indications of speech-rhythmicity, or at least the absence of song- or music-rhythmicity. The phrases ‘Roses in the road / teddy bear outside / bullet there on the right’ are clearly parallel rhythmically and thematically (and even melodically), but because the rhythmic figure is performed three times, not two or four, it contributes to the overall irregularity of the verse. This small section also destabilises the music-rhythmicity in another way at the quantised zoom level, as can be seen in figure B4.4.

![Figure B4.4: Nested duple cross-rhythm in Noname’s ‘Casket Pretty’ (from 0:53).](image)

Whereas Noname uses a quite regular triplet subdivision in the beginning of the verse, which could be characterised as ‘dactylic metre’ in poetic analysis or as a ‘triplet flow’ flow type following Duinker (2019), there are occasions, as in figure B4.4, where she introduces a cross-rhythm with stresses on every other syllable rather than every third. This creates sections of three over two which feature three groups of two over two groups of three. However, here the microrhythmic disalignment from both the backing track and a relatively regular triplet subdivision is severe enough that this cross-rhythm is not experienced like typical cross-rhythms in a groove. This disalignment, that is, does not fit snugly onto the existing rhythmic and metrical framework to enhance or comment on the existing groove by introducing a complementary but slightly destabilising alternative. Instead, it performs the same role as the stuttering irregular musical metre—it tries to break free of the repetitive regularity and symmetry of the established metrical frameworks. It introduces a competing rhythmic layer, but one that is unstable and uneven. So, rather than creating an added level of
rhythmic complexity by nudging at the stability of a dominant framework, this cross-rhythm attempts to completely disrupt the framework upon which it is superimposed.

While most of these rhythmic techniques and features are not necessarily reminiscent of actual speech, they do display some of the features that we have established as typical of speech-rhythmicity. First, the irregularity at every rhythmic zoom level is atypical of music-rhythmicity. Second, the shape of these irregularities also follows what we would expect from speech-rhythmicity, in that they tend to start out more or less regular and become more irregular towards the end of phrases/formal units. This is evident in the musical (hyper-) metre, where the one bar of differing length arrives at the end of the pattern. Also, many of the long phrases (uninterrupted by pauses) display instability towards the end by not rhyming in the expected manner or ending on the expected (music-) metrical positions. All in all, the verse is so actively not song-like rhythmically that, if we accept that there is a single axis of rhythmicity from speech-rhythmic to music-rhythmic, this verse approaches the pole of speech-rhythmicity. Yet, when we listen to the track, it is clear that Noname’s delivery is not particularly speech-like, mostly because of the way in which she traverses the other in-between-space of speech and song: the intonation.

If Noname’s rhythmicity is speech-like through being not song-like, her intonation is probably best described as song-like through being not speech-like. She rarely fully breaks into song in the sense of long extended vowels or melismata, but her melodies both centre upon discrete pitches which match the tonality of the musical background and employ both a range and a local variation (large melodic leaps within phrases) which are uncommon in speech. A simple ear test should make it clear that Noname’s intonation is something different from both most rap and common speech. The adjective that comes to mind is ‘sing-songy’, as the melodic contours display a song-like range and she rarely sticks to an extended M-tone (cf. Hognestad’s analyses), choosing instead to change the musical pitch on nearly every syllable, as can be seen in figure B4.5.

**Figure B4.5**: Melodic contour of the phrase ‘Roses in the road, teddy bear outside, bullet there on the right’ (from 0:53) in Noname’s ‘Casket Pretty’, extracted using Celemony’s Melodyne software (version 4.2.4).
Figure B4.5 shows the melody of the phrase also analysed in figure B4.4. As Noname repeats the rhythmic figure three times, she repeats the melody as well, save for the (slightly sharp) Ab in the first instance of the phrase being raised to (a slightly flat) A in the two repetitions, matching the chord change from Ebmaj7 to C/D (or D9sus4, depending on notational preference). This melodic phrasing is not speech-like at all and is perhaps less speech-like than most of the rest of the verse. Noname may have chosen this specific section for a more consistent stretch of singing to strengthen the parallelism between the repetitions of the figure, emphasising the cross-rhythmic destabilisation of the rhythmic framework.

Of course, that explanation is an interpretive stretch, as one does not need a rhythmic or formal reason to phrase a melody in an interesting manner, as Noname demonstrates with the phrase in figure B4.6.

**Figure B4.6**: Melodic contour of the phrase ‘Somebody hold me, somebody take my hand’ (from 1:20) in Noname’s ‘Casket Pretty’ (Melodyne version 4.2.4).

Here, the melody spans a full fifth, but there is still an inexactness to the specific pitches (most notably in the two huge glissandi in the first two syllables) which serves to restrain the song-likeness of the melody. Again, Noname’s delivery is more un-speech-like than fully sung, operating in the in-between-space between speech and song’s respective tonalities.

Considering both rhythmicity and tonality, Noname’s idiosyncratic style\(^{165}\) of utilising the in-between (of speech and song) raises some interesting points. First of all, while there are axes of speech- to song-likeness within the dimensions of both intonation and rhythm, these axes seem perfectly capable of operating independently from one another. While the analysed

---

\(^{165}\) Note that ‘Casket Pretty’ is perhaps the most extreme example of Noname using this not-quite-singing sing-songy intonation. In most other tracks, she trends towards more ‘fully sung’ intonation with more use of repeated notes, like the style employed by Bone Thugs’n’Harmony described earlier in the track.
passages from ‘Casket Pretty’ are probably the most ‘speech-rhythmic’ of all the musical examples presented in this thesis, the intonational phrasing is very un-speech-like. This ties into the second point: ‘Casket Pretty’ is an example of an uncommon style of singing which displays some speech-like intonational features (the extensive use of glissandi, the use of unstable/imprecise musical pitches and the lack of melismata) but in a very exaggerated way which makes it clearly different from common speech. However, its intonation style is also clearly different from most rap (quite large and varied melodic range within phrases, no stable M-tone and so on). This points to yet another complication if one is trying to define ‘rap’ utilising the spectrum from speech to song: there are so many different ways to manipulate both rhythm and melody within this spectrum which can be either rap typical or rap atypical.

Is ‘Casket Pretty’ rapped or sung? Or both or neither? Is Noname a singer or a rapper, then? Both? Something else?

**Utilising the in-between 2: ‘SONGS ABOUT YOU’**

While Noname’s ‘Casket Pretty’ is a study in nuance and degree of in-between-ness, raising questions about the boundaries between rap and song and what speech-rhythmicity entails at different rhythmic zoom-levels, Chika’s ‘SONGS ABOUT YOU’ (2020) traverses the in-between-space in a very different manner. Showcasing a much more ‘music-rhythmic’ microrhythmic profile, Chika’s manipulation of speech-to-song-likeness derives from the way in which she sections parts of the song’s form and parts within the song’s formal units by switching between a fully sung and ‘fully rapped’ (if such a thing can even be defined) intonational delivery, both gradually and abruptly. In addition, the use of backtracks/dub tracks by herself and her backing singers adds an additional layer of in-between-ness on occasion, where rap and song are both present simultaneously.

The opening passage of Chika’s first verse, transcribed in figure B4.7, is a good showcase of these techniques.
Figure B4.7: Transcription of the first six bars of the first verse in Chika’s ‘SONGS ABOUT YOU’ (from 0:23).

Rhythmically, the flow is a variation upon the common sixteenth-note-stream flow type, and Chika’s swung sixteenth notes sit well in the groove, driving the track forward. The occasional stressed syllable in an off-beat position introduces some variation to the stream, along with the increase in rhyme density in bars 3 and 5 and the variation in the length of the syntactical units of the lyrics. However, the star of the flow’s show is the use of various intonation techniques.

The verse starts with a conventional rap intonation oriented around an M-tone of Ab but without fully committing to using only discrete musical pitches (I choose to transcribe passages like this using cross noteheads and not indicating any relative pitch movement). In bar 3, however, the intonation gradually switches to fully sung. The change in Chika’s intonational delivery is relatively small, but with the support of the backing vocals’ doubling of the words ‘Alabama bitch. Fuck getting rich’, the ever so slight pitch quantisation166 which

166 By this, I am referring only to Chika’s vocal delivery, not to any production effects. Here, that is, Chika switches from approximate musical pitches to precise pitches—closer to canonical digital pitch values—in a manner resembling a quantisation of the pitch classes.
Chika inserts into her melody creates something of a new gestalt. The flow is now no longer ‘pitchless’ (or, more precisely, displaying speech-like, not song-like, intonation). It instead boasts a clearly defined musical melody, even if the actual measurable pitch content has changed very little.\(^{167}\) However, the intonational mode which Chika has entered is still of a particular kind of ‘sung rap’, as there is a minimal pitch range of two notes separated by a minor third (Ab and F, emphasising the F minor tonality of the musical background), and Chika pivots between them throughout bars 3 and 4. When she breaks out of the sixteenth-note stream via her off-beat phrasing in bar 5 (with a syncopated pickup), this rhythmic change is accompanied by an increased range and note selection in her melodic phrasing. This shifts the mode of delivery from a kind of ‘sung rap’ to undeniably ‘fully sung’. Note, however, that there is a clear parallelism in the formal structure between the less-sung bars 3–4 and the fully sung bars 5–6. The primary rhyme is revealed in short syntactical segments in the first bar of the line pairs (with ‘bitch’ and ‘deliver’, respectively), before a longer line without an internal rhyme instance signals the end of the short rhyme complexes (with ‘disc’ and ‘with her’). Thus, Chika’s switches in intonational mode do not interrupt the flow’s continuity but rather serve as another means of building up the track’s formal structure, the entirety of which is visualised in figure B4.8.

---

**Figure B4.8:** Song structure of Chika’s ‘SONGS ABOUT YOU’.

\(^{167}\) This is reminiscent of Diana Deutsch’s famous ‘speech-to-song-illusion’ (see, for example, Deutsch, Lapidis, and Henthorn [2008]), wherein speech easily morphs into song even when it is not repetition which causes the perceptual shift (as is the case here). Both Komaniecchi (2019, p. 116) and Ohriner (2019a, p. 11) discuss this floating boundary between speech and song in the context of rap flows.
Chika builds up the formal song structure of ‘SONGS ABOUT YOU’ through vocal performance, arrangement and production more than through the organisation of the musical background. While there are clear instrumentation and arrangement variations in the looped eight-bar chord progression (as indicated in figure B4.8), it is the vocals which tie the different sections together across the larger formal units of ‘verses’ and ‘choruses’ and also clearly mark the difference between the fully sung chorus and bridge and the verses.

Interestingly, the way in which these larger units are organised is relatively uncommon. The section dubbed ‘bridge’ in the figure is not a typical bridge. Because it shows up early in the track, after the first chorus, one would expect that it is either the chorus, with the previous section being a ‘pre-chorus’, or that it is part of an especially long sixteen-bar chorus. However, the section only appears once, so I find the label ‘bridge’ to be serviceable. The verses are also unconventional, in that they vary in length: the first spans a single eight-bar unit, the second spans three of them (for a total of twenty-four bars) and the third spans two (for a ‘standard’ sixteen bars). Additionally, the song ends with a verse rather than a chorus, which is by no means unheard of in a rap context but is definitely not the most common way of organising a (pop-) song structure.

Between the verses, there are parallels as to how the various eight-bar symmetrical blocks are structured. The first verse sets a blueprint which is almost-but-not-quite repeated in the middle block of verse 2 and the first block of verse 3. These blocks, marked with the letter A in figure B4.8, begin with a rapped delivery, veer into song after the halfway point of the block—further subdividing it into two four-bar units—and return to the rapped delivery again. Interestingly, the way this is accomplished is slightly different in each iteration of A. Whereas the first verse features the aforementioned gradual categorical morph from rap to song, the second and third verses display clear categorical boundaries between rapped and sung modes of delivery. It is as though there has been an embrace of the gestalt switch from speech-intonation to song-intonation, and it is now impossible to blur the categorical boundary in the same manner again (except perhaps in the passage from 3:44). Instead of being a feature of Chika’s delivery, the mixing of rap and song in verses 2 and 3 is accomplished through the use of sung backing, doubling Chika’s words and flow with a fully sung and harmonised melody. This arrangement technique is used on the second and fourth lines of the blocks, again creating and emphasising symmetrical units subdividing the eight-bar block. At the fifth bar of the second A section (the middle section of the long second verse), Chika switches to a fully sung delivery again, perfectly mimicking the structure of the first A section. However,
rather than gradually constraining the melodic range through bar 7 before returning to a fully rapped delivery as was the case in verse 1, this block is more symmetrical, and the delivery switch happens at the start of the seventh bar.

This second version of A represents a formalisation of the ideas from the first verse, tightening them up and making the section symmetrical with clear categorical boundaries between the various intonational techniques. The third version of A (the first part of the third verse) goes in the opposite direction by taking the symmetrical blueprint and playing with its established expectations. The sung doublings of the second and fourth lines are preserved, but rather than repeating the fully symmetrical convergent metrical structure from the first and second versions, Chika plays with a divergent metrical structure by increasing the rhyme density and placing the line endings with the sung doublings on the third beat of bar 2 and as a syncopated one-rhyme to bar 5. The main structural reference points are preserved, but the symmetrical idealised version is challenged. Another parallel is broken in bars 5 and 6 as well, as Chika does not switch to the expected sung delivery but instead leaves the singing to her backup singers, who add an extra sung doubling before a huge crescendo on belted ‘oh’-s builds up a cadence to the end of the section (which is succeeded by a final rapped B section rather than the perhaps more obvious chorus ending).

Chika’s use of different techniques to emphasise the symmetrical sectioning of the rhythmic structure and create parallels across the formal sections is a way of employing the in-between which is less about categorical ambiguity (besides the passage in the first verse) and more about using categorical difference along an axis which is perhaps not the most utilised within a single piece of music: that of melodic quality and layering. While ‘Rappers can and do use several ( . . . ) pitch techniques simultaneously or in close proximity in a single track or verse’ (Komaniecki, 2019, p. 118), the techniques employed in ‘SONGS ABOUT YOU’ range from the most speech-like to the most song-like (to co-opt Komaniecki’s terminology), and they do so in a manner which is unusually systematic. The techniques become not only potentially expressive and aesthetic effects in themselves but also tools to create and/or emphasise the larger-scale formal structure—another layer of aesthetic expression. And while Chika uses techniques like those described by Komaniecki which are features of a single vocal track, some of the most significant pitch techniques in the flow of ‘SONGS ABOUT YOU’ are not in fact part of the main vocal track.

This (re-)raises the question of what the actual constituents of a rap flow are. A rap flow as a composite auditory stream is not necessarily only one voice track with its inherent
accompanying reference structures (including syntactical and phonological boundaries, various parallelisms, patterns of stressed syllables, and so on). It can be, and most often is, a tight composite of layered vocal tracks. These tracks can range from being almost inseparable from the ‘main track’ (often referred to as ‘dub tracks’) to being clearly, even overtly separate. In track B2, note 134 (page 179), I briefly discussed Kendrick Lamar’s use of a gradual balance shift between the main track and the effects-processed ‘backtracks’ in ‘F*ck Your Ethnicity’, where the composite layers of the flow stream are pried gently apart, revealing the individual layers and shifting the focus from one layer to another (or others). The techniques Chika employs on ‘SONGS ABOUT YOU’ are not about gradual shifts but rather about abrupt reconstructions of the combinations of layers in the composite stream. In both cases, however, the ‘rap flow’ does not consist of a recording of one single vocal performance, or even the illusion of one (via perfect comping and blended dub tracks). Rather, these are examples of a phenomenon described by Di Bona (2019) as chimericity.\footnote{Di Bona employed the term on passages of classical music. The appropriation of the term to rap flows is my own.} Adopting Bregman’s evocation of the mythological chimera—where a single (composite) auditory stream consists of several, clearly separate constituents (Bregman, 1990, p. 459)—Di Bona explored different musical applications of chimericity (‘melodic chimeras’, ‘rhythmic chimeras’ and so on). The quality of chimericity involves the fact that one hears and experiences both a musical ‘line’ as well as the shifts in which instrument or combination of instruments expresses that line. Strict musical chimeras will jump completely from instrument to instrument (the ‘Klangfarbemelodie’ technique popularised by Schönberg and Webern is a good example), while lines where there is shift in the composition of or balance between different instruments will express a gentler kind of chimericity. The blending of different vocal tracks in rap flows is an intuitive parallel to the many heads of a chimera.

Chika expresses several different degrees of chimericity in ‘SONGS ABOUT YOU’, but there is always a main lead vocal track around which the composite flow is centred. The simplest and most typical gestures are of the kind heard in the first B section, the beginning of the second verse (from 1:35), where the flow is carried by a main track and a dub track that doubles parts of it. The chimericity in passages like this (which are often whole verses or tracks, where there is only ever a single main track and one or more dub tracks) resides in the contrast between the parts which are doubled and those which are not—they are the same, but they are also different, even if only slightly so. The end of the fourth bar introduces a different
As discussed, Chika uses the layering of tracks to morph from rapped to sung delivery in the first verse, where the introduction of sung doubling works to introduce (musical) pitch stability to the rapped main track before that track embraces the sung mode completely. There are variations to this layering of rap and/or song throughout ‘SONGS ABOUT YOU’ which create both slightly different aesthetic effects and varying degrees of chimericity. There is the rapped main track with a monophone sung doubling (which is only really used during the morphing passage in verse 1); there is the rapped main track with harmonised doublings (such as lines 2 and 4 in the second and third A sections), and there are the combinations of sung main track and sung backing tracks (and even the ways in which the sung doublings omit words or parts of phrases). As seen in figure B4.8 above, these techniques are typically applied at specific points within the form. The ogre voice, for example, marks the ends of four-bar blocks, typically only emphasising a few words towards the ends of lines. The rapped main plus sung doublings parts create the previously discussed parallelisms between the verses. In passages with sung main track and sung backing, the sung backing tends to only

169 ‘Ogre’ is a colloquialism employed for this technique among those involved in my group’s label.
170 There can also be variation between the different instances where there is an added ogre voice in terms of its prominence—that is, how loudly it is mixed.
171 On the track, it sounds as though these harmonies are sung by Chika herself and then extensively tuned and tightened up in post-production. There are other backup singers on the choruses and the wordless backing vocals in the verses who sound like the same backup singers who perform with Chika live, where they also sing the sung doublings. This raises an interesting question of the difference between composite flows consisting of different tracks/voices of the same person’s voice versus the same musical content performed by other people (such as the traditional ‘hype man’ or, as in this example, backing singers). Taken further, it raises questions about the impact of timbral qualities in voices and their ‘blending’ and the difference between stacked layers of a single singer, ‘family groups’ of very similar voices (think of the Beach Boys or the Everly Brothers), and more dissimilar voices.
172 In the passage starting at 2:17, the ogre-doubling stretches over two full bars at the end of an eight-bar block (the second A section), but the volume level of the doubling is very low on everything except the final word ‘medicine’ (where there is also added distortion).
double some words rather than entire phrases (in contrast with the way in which the sung
doubling of the rapped main track is employed). Even though there is a main lead track which
remains a thick common thread in the middle of this tapestry, the chimericity evoked by the
variation between the different types of layerings is a significant part of the experience of
‘SONGS ABOUT YOU’.

Particularly applicable to the overarching question of how artists manoeuvre the in-between
are these layerings consisting of rapped main tracks and sung backing. When assessing how
speech-like or song-like a passage of a rap flow is, the parameters that Ohriner, Komaniecki,
Hognestad and others analyse tend to focus on a single voice and the ways in which that
single voice expresses specific speech-like or song-like features. When we are considering the
rap flow as a composite auditory stream, however, it can become a chimera with (at least) one
speaking head and another singing head. Of course, one can still place such flows on a
spectrum from speech (or rap) to song, but it is hard to compare a blend of different speech-
and songlike tracks to a single such track. How much more should one weight the ‘lead track’
as opposed to a backing track? Would a perfect blend of a fully spoken and fully sung track
be 50 percent song-like in all? At the very least, it seems clear that the use of layered tracks
and chimericity is yet another way for rappers to activate the in-between, and that being not-
quite-song and not-really-speech is yet another of rap’s essential aesthetic qualities which is
rooted in categorical ambiguity.

While rap clearly cannot be defined simply through its melodic and rhythmic parameters as
occupying some space between speech and song—as exemplified by Bizzy Bone of Bone
Thugs-N-Harmony (figure B4.1), something can be defined and categorised as ‘rap’ while
still being clearly and unequivocally ‘sung’—there are multiple techniques and approaches
rappers can utilise which, in fact, originate in this in-between-space. A rap flow, or parts of it,
can be more or less ‘speech-like’ along both the rhythmic and the melodic axes, but it can be
so to different degrees and in different ways, and it does not need to manoeuvre along both (or
either) axes either. Noname, for example, uses a speech-like rhythm, while her melodies are
more song-like (albeit in a quite atypical and idiosyncratic way), whereas Chika has a less
speech-like rhythmic phrasing but uses a rap-typical constrained melodic range in her
‘rapped’ sections (she, of course, has clearly contrasting sung sections with greater melodic
ranges and a clearer emphasis on musical pitches). It might be impossible to pinpoint the
exact boundary between song and rap, but there can be no doubt that rappers have, in general,
a very sophisticated understanding of the rhythmic and melodic possibilities afforded by the
gradated and ambiguous landscape between speech and song.
Track B5: What makes the shit dope?

Side B has been dedicated to an exploration of various techniques rappers employ in their flows. In track B1, the techniques revolved around lineation and bridge rhymes and different types of heteromorphies where the relationships between bars and the boundaries of rhyme complexes create structural tensions and asymmetries. Similarly, track B2 kept the lens affixed to the macrorhythmic zoom level, exploring how rappers use the gravity of the form to structure their flows. Track B3 zoomed in on the quantised and microrhythmic zoom levels, showcasing techniques where microrhythmic disalignments between various layers of the composite auditory stream can result in categorical ambiguities. Finally, track B4 delved into the melodic dimension of rap to show how rappers utilise the in-between space between song and speech for various rhythmic effects.

Common to all these techniques is a theme that has been central to the thesis as a whole: the interesting things (or dope shit) are found around points or areas of convergence—that is, places where concepts, features, structural parameters or ideas collide and intermingle. Of particular interest are the many ambiguities and dualities which arise when boundaries overlap or blur together. In this final track, then, I will revisit and summarise the different convergences, ambiguities and dualities which set the stage for this thesis’s exploration of rap flows before turning to those which comprise my answer to my initial research question: What are the core rhythmic features of rap flows? Then, I will discuss my approach and answer to the second research question: How are rap flows analysed? That is, how were the answers to the previous question produced, and how do these answers vary depending on the analytical framework in use? Specifically, I will contextualise how and why the approach I have chosen is suited to illuminating what’s dope in rap flows, and how it can be adapted to and further developed for different directions of future scholarship and practice.

Convergences, ambiguities, dualities

Invoking the proverbial Ouroboros and returning to the starting point of this thesis, there are certain points of convergence which have been foundational to every track. My perspective as a scholar/rapper/enthusiast is a congregation of different potential angles from which to investigate rap flows, with all the pros and cons that this brings with it. I certainly have many biases and blind spots in my approach stemming from my emotional connection to the music and the habits and preferences imparted by years of practical engagement with it. For example, my choice of analytical examples is not based on any objective unbiased selection.
(like the corpus studies of, among others, Condit-Schultz or Ohriner presented in track A2) but instead on my familiarity with them. Nor are they chosen to highlight features which are identified through statistical analysis or through what a large number of fans identify as important, but rather through what I believe to be important. This approach, my biases, my background and my positioning have to be taken into account when engaging with my analyses. However, the specific positioning of a rapper/scholar is not available to everyone, and it has not been exploited much in the field to date, even though it clearly brings with it a potential for new insights. As Southern stresses: ‘Serious study of African-American music requires getting to know the music, which means listening to it and, if possible, performing it’ (1997, p. xx). One intersection which emphasises this point is that between musical analysis and musical practice, as presented by Agawu (2004) and discussed in track A1. As the act of analysis is (or at the very least can be) in itself performative and compositional, embracing the individuality afforded by the analyst’s identity can enhance the value of the analysis. In the analysis of rap flows specifically, an angle such as this is even more appropriate—rather than merely doing ‘analysis of hip-hop’, one can take inspiration from the hip-hop ethos and do ‘hip-hop analysis’, and, as Rollefson (2017, p. 10) says, ‘take on the mantle of “hip hop scholar” by doing scholarly work in a hip-hop way’.

For me specifically, a duality that is important to acknowledge and reflect upon is the contrast between the Blackness of the cultural practice of hip-hop and the Whiteness of my heritage and background. My goal has been to make sure that when Black and White blend, the result is not grey but rather zebra-striped—the Blackness shall be undiluted, but the Whiteness is not hidden. My hip-hop scholarship is obviously from a version of the culture which has itself been imported and altered—a ‘Scandinavian fusion’, if you will. As I discussed at length in track A1, the global appropriation of hip-hop culture brings with it a multitude of ethical issues. For example, since hip-hop has completely permeated the cultural mainstream and been thoroughly commercialised, there are many more people who engage with, adopt, adapt and appropriate parts of hip-hop culture—and it is not a given that all of them will treat hip-hop in a culturally sensitive way. I am well aware of my position of privilege, and I strive to use that in a socially and culturally progressive manner because those of us who have chosen to engage with hip-hop culture (or other Black cultural expressions) in the academy have a special responsibility to engage with it conscientiously. I hope and believe that I have managed to preserve the Black essence of rap flows as a musical constituent (in all of the examples from all of the different artists—whichever cultural background they might have),
and that my perspective as an outsider-insider (that is, a Norwegian rapper) can be illuminating to readers of any nationality, colour, creed, cultural belonging or disciplinary or scholarly background.

To take the spotlight away from me—the analyst—and return it to the truly significant star of the thesis—the music—there are several important epistemological dualities which have been central to this thesis. First, the understanding of rhythm as being an amalgam of an auditory signal and the listener’s perception of that signal—of ‘non-sounding reference structures’ and ‘sounding rhythmic events’ (Danielsen, 2010a, p. 4), or of figure and gesture. The analyses in this thesis have focused on how a listener (both myself and a hypothetical alternative listener) experiences the rhythmic structure of the flow, not the ‘objective truths’ of the measured acoustic signal. However, the measured signal was nevertheless used as a reference (rhythm arises in the meeting of the sound signal and perception, after all), and different types of transcription and representation do tend to present themselves as ‘objective’ when they are formalised, even though they always remain some sort of perceptual interpretation. (Also, the measurement of microrhythm using waveforms and spectrograms serves as a foundation of my own personal interpretation.) Thus, my focus on ambiguities, alternative interpretations and the process of constant reinterpretation as the music unfolds has been, in part, to emphasise these conditions of analysis. The analyses themselves are windows into my listenings, above all else—others’ may well be different.

Furthermore, the intersection between the digital and the analogue dimensions of rhythm perception, as presented by Kvifte (2004) and discussed in track A3, is at the heart of many of this thesis’s theoretical and analytical results. The discrete categorical belongings of rhythmic events (rhythmic structure) and the positions of these events on a continuous scale within categories (rhythmic expression) comprise the dual nature of our perception of rhythm—they are perceived simultaneously, and they interact with and inform one another over time. I will return to exactly how this nature is central to the answers to my research questions, but in general terms, the boundary between what is structural and what is expressive is a key feature of many microrhythmic analyses (as discussed in track A6 and explored via analysis in track B3). Also key is the way in which rhythmic expression tugs at and challenges the categorical clarity of rhythmic structure. It is important to stress, however, that digital/analogue and structural/expressive are not synonymous pairings even though they may appear to be (see tracks A1, A3 and A6). Not all analogue modulation is expressive, for example—it may be systematic and structural. An analogue feature which might be experienced as expressive in
isolation can become structural through repetition and/or its relationship to other rhythmic events.

This floating boundary also relates to rap’s constituents, as derived from Edwards (2009): content, delivery and the main research object of the thesis, flow. While the taxonomic division seems neat and intuitive, and it is an integral part of the discourse around rap, it is not necessarily easy to disentangle the categorical boundaries, particularly between flow and delivery, as is evident from the extensive discussion of the term ‘flow’ in track A1. As discussed in tracks A1 and A6, I consider the relationship between flow and delivery to be like the one between structure and expression, but mine is not the only definition. Some define articulative features, which I consider mainly expressive and part of delivery, as part of flow (systematic articulative features, of course, can be structural and thus part of my definition of flow). Similarly, some consider rap’s melodic features to be part of flow, whereas I delimit flow to rhythmic structure and consider melodic structure to be a different parameter altogether because it does so much more than impact rhythmic structure alone.

This leads to a duality which was explored at length in the previous track (B4)—namely, that of song and speech. My separation of melody and rhythm in my definition of flow is not to devalue melody’s importance to rap but the opposite. Rap’s melodies, as mentioned, are much more than just a part of the rhythmic expression, and this insight is informing exciting new developments and publications in the field of rap analysis. My analyses, on the other hand, were principally focused on flow and musical rhythm, and especially the ways in which rappers use melodic features to create rhythmic effects which are both structural (including parallelisms across large formal units) and expressive. Both speech and song encompass rhythmic and melodic features, and while I maintain that rap should be considered a type of song (and not a type of spoken word), rappers use features of speech-likeness extensively, venturing much closer to speech along the rhythmic and melodic axes between speech and song than is common in most types of vocal music. Again, it is in the between-space that dopeness appears.

These are not the only convergences, ambiguities and dualities explored in this thesis, of course, but not everything needs to be reiterated and summarised. I will now turn to the goal initially formulated in the introduction—the answering of the research questions.
Research question 1: What are the core rhythmic features of rap flows?

This first research question is a big one—only slightly less grandiose than the titular ‘What makes the shit dope?’ To make it more manageable, I divided it into three sub-questions which gradually hone in on the rhythmic language of rap flows:

What characterises the structural framework(s) of a rap flow?

How do rap flows interact with and contribute to creating their inherent structural framework(s)?

What are the essential rhythmic techniques—structural and expressive—which rappers employ in this interaction?

While much of the answer to the first sub-question is found in the existing literature—in particular, the various corpus studies presented in track A2 and their data-driven extraction of global features—there is an important nuance in the difference between ‘rap track’ (or ‘rap song’ if one is extending the analysis to live performances as well as recorded music) and ‘rap flow’ with regard to their structural frameworks. However, some broad generalisations can be made concerning song form and verse structures which are held in common by the two.

First, almost all rap music is in some form of 4/4. There are exceptions, like the example of Noname’s ‘Casket Pretty’ analysed in track B4, but the vast majority of rap tracks conforms to the ubiquitous quadruple metre. As Gudnason said in his article discussing how rap lyrics take shape on paper, it is ‘a form that in some way has to have a relationship to the number four’ (Gudnason, 2016; see also track A4, page 70). For the rapper, this is true even in the rare cases where there is another musical metrical framework. By deviating from the common 4/4, one has to alter the approach one normally takes when coming up with (or ‘composing’, even if most rappers do not use the term) the flow. The rapper’s habits and rhythmic toolbox are tuned in on 4/4, and if one is devising flows to suit (or create) uncommon metrical frameworks, this ‘normal’ condition and how one chooses to deviate from it inevitably becomes a reference point.

Second, this 4/4 musical metrical framework will typically be organised into symmetrical blocks. The four-bar block is perhaps the most ubiquitous, with its doubling—the eight-bar unit—being the standard length of a chorus or ‘hook’. Slightly less universal but still clearly a ‘standard’ is the common verse length of sixteen bars. The hegemony of this quadruple symmetry of four blocks of four bars has been challenged somewhat by the rise of the twelve-
bar verse. While the eight-bar period is a strong symmetrical unit in a sixteen-bar verse, this is not really the case in a twelve-bar verse, as it no longer subdivides the verse into two equal parts. The typically looped basic units of the instrumental track in a rap track are more likely to be four bars than eight bars as well, meaning that the four-bar unit is the most important symmetrical building block in the genre as a whole.

At this point, however, the differences between ‘rap track’ in general and ‘rap flow’ begin to appear. The relationship between musical background and flow is another essential point of convergence—one which is very relevant to my second sub-question: How do rap flows interact with and contribute to creating their inherent structural framework(s)? For one thing, the quadruple grouping of bars is stronger than the duple grouping not only because the instrumental track suggest groups of four bars more often than groups of two but also because of the ways in which rappers organise their flows. One characteristic of rap’s common structural framework which is central to the flow’s interaction with and contribution to that framework is the convention of the line/bar coincidence. While the most common way of organising lines and bars is to have them coincide, rappers will routinely create tensions among flow and beat and poetic lineation and musical metre by disaligning the boundaries of lines and bars. This means that there is room for variation within the smaller metrical groupings.

In track A4, I introduced the terms convergent and divergent metrical structure, where the former denotes adherence to the line/bar coincidence and the latter denotes sections of flow where lines and bars do not coincide. To briefly reiterate, the realisation of ‘metre’ and metricality varies a bit in music and literary theory, respectively. The levels of metrical units, subdivisions and spans—groupings of units—correspond to one another but do not necessarily coincide. Whereas musical metre has beats which are grouped into bars, poetic metre’s beat-level units are called ‘feet’, and the spans which correspond to musical bars are called poetic lines. The act of dividing a text into lines is called lineation, and as I have argued—in agreement with Fabb (2015)—this process generates an implied form, meaning that it is dependent on the triangulation of various types of evidence for lineation. There are many types of such evidence, some stronger than others, and the ones upon which I have focused most in this thesis are rhyme and syntax.

---

173 This can be seen in the graph on page 144 of Condit-Schultz (2016c), where the twelve-bar verse gradually becomes as common as the sixteen-bar verse starting in the late 2000s. This trend appears to have continued in the years since Condit-Schultz’s study was published.
Through convergent and divergent metrical structure, rap flows *interact* with the structural framework both at the basic metrical level and at the level of larger symmetrical blocks. Additionally, as alluded to when I identified the four-bar unit as stronger than the smaller two-bar unit, the flow *contributes to creating* the structural framework. When applying a divergent metrical structure, rappers will often emphasise sections of four bars while disrupting symmetry at the two-bar period level. A prime example is Lars Vaular’s ‘Helt om natten, helt om dagen’ in track B2, but this approach is also prominent in figure A4.7 (page 80) and figure B1.10 (page 148), among others. In this way, rappers can strike a balance between reinforcing and emphasising the expected structural framework at one level while breaking expectations by deviating from the expected convergent metrical structure within that structural block.

There are many examples of flows which can be understood as simultaneously interacting with and contributing to create their structural framework(s), and most of them do so by creating some sort of structural ambiguity. One approach is to use lineation to create repeated periodicities which ‘compete’ with the musical metre. Again, returning to the example from figure A4.7, the four bars of Dizzee Rascal’s flow on ‘Dirtee Cash’ are structured with lines spanning three beats each, clearly evoking a competing 3/4 metrical framework alongside the relentless four-on-the-floor grime beat. Similarly, in passages of *repeated one-rhyming* (see track B1), the flow presents a displaced 4/4 musical metre. Thus, there is a clear reinforcement of the general ‘4/4-ness’ of the structural framework but still a marked destabilisation of the expected convergent metrical structure. Other times, the flow might emphasise the halfway point of the musical metre, either through displacement—as in ‘Anaconda’ (track A4, page 68) and ‘F*ck Your Ethnicity’ (track B2, page 178)—or by subdividing the musical metre either completely (through short lines in slow tempi) or through weak alternative lineation (some lineation evidence indicating half-line phrasing but falling short of overpowering the dominant lineation).

This last example is related to how the flow is part of determining which tempo is the dominant one in *dual-planed* rap metres (discussed in track A6). In cases where there are two reasonable interpretations of what the dominant tempo might be, the flow can decide which is most salient to structuring the listening experience. In the ‘trap’ styles used as examples in

Note the similarity with Ohriner’s (2016a) analysis of the way in which T-Mo Goodie emphasises a triple metre in his phrasing on OutKast’s ‘Mainstream’, in contrast to the track’s other verses. The difference between ‘Mainstream’ and ‘Dirtee Cash’, however, is that in the former the instrumental track invites both metrical interpretations, whereas the flow creates the cross-metre all on its own in the latter.
this thesis, there is typically a slow boom-bap pattern with a density referent of sixteenth-note triplets. Depending on how slow the boom-bap is, the double-time interpretation might feel more natural. The flow, then, has a key role, as the lineation (through rhyme position and syntax) might reinforce or contradict the double-time feel. The example of ‘I Like It’ (discussed in track B1, page 139-140) reflects this duality well, as the flow switches between emphasising the different tempi (listen from 1:24). It is crucial to note that while I transcribe such flows with the slow boom-bap tempo as the dominant one, the constant negotiation between the two tempo planes is a significant part of the identity of these types of metres. It is not a case of either/or, nor do we necessarily constantly shift between two competing metrical frameworks. Instead, both are present at all times, and the ways in which the various musical constituents sometimes reinforce one or the other is part of the aesthetic ambiguity which makes the metre what it is.

The examples of flow interacting with and contributing to creating its structural framework(s) have thus far been at the span-level of musical metre and poetic lines. The meso-level (my quantised zoom level) is also home to several important aspects of this interaction and creation. Syncopation, crossrhythmic and counterrhythmic phrasing and the like will all influence the structure of our rhythmic experience. As previously discussed, these types of features are explored extensively in some other key works of rap analysis, like Ohriner’s ‘Flow’ (2019b) and Komaniecki’s dissertation (2019). Less explored is how microrhythm influences the structural framework(s) of a flow.

While it is intuitive to couple microrhythm with expression rather than structure, there are several ways microrhythmic features of flow can impact the structural framework(s). At a general level, an extensive use of expressive timing (or, more precisely, microrhythmic phrasing which is disaligned from canonical time-value classes) will tend to make the entire rhythmic composition of a rap track feel ‘looser’, in a sense. Microrhythmic variation makes rhythmic structure less clear cut, as it introduces ambiguity and invites alternative structural interpretations. If there are multiple asynchronies between rhythmic events in different layers of the composite auditory stream, it increases the chance of non-simultaneous events being interpreted as belonging to the same category—that is, it widens the beat-bins. A syllable that is ‘late’ in comparison to the beat position presented by the drum tracks might be experienced as part of the beat in a context where the timing is ‘loose’ (due either to a consistent slight disalignment of repeated elements or a general use of expressive timing), while if the timing is ‘tight’, a similarly late syllable might be experienced as belonging to a different rhythmic
category (not part of the beat). This relation is heavily influenced by how much categorical rhythmic information is present, and how clear it is; the less information there is in the instrumental track, the more influential the flow (and the delivery) will be. In the passage from ‘Si eiga rås’ (figure B3.2, page 187), for example, most of the elements of the musical background are removed, so the flow alone dictates the rhythmic structure and metrical framework (in conjunction with the established reference structures, of course). In the examples from ‘Eple’ (figure B3.4, page 194) and ‘Do You Remember’ (figure B3.5 page 197), the musical backgrounds are also rather sparse, or at least not very insistent upon pinpointing the exact locations of some rhythmic categories, and this too gives the flows, and the way they are phrased, significant structuring power.

Some rhythmic events are also more perceptually significant than others. The concept of metrical anchors entails the expected convergence of certain stressed syllables and metrically strong beat positions in the musical metre. Even if there is a significant disalignment between the metrical anchor in the flow and the structuring elements of the musical background, the metrical anchor will strongly communicate a categorical identity. For example, in the passage from ‘Waterfalls’ in figure B3.1 (page 184), the temporal position of the metrical anchor is around 190 milliseconds later than the musical background’s communicated beat position—a distance equivalent to a long swung sixteenth note (that is, the first of the two uneven sixteenth notes in a swung subdivision) in the relevant tempo. This does not mean that there is a consistent huge beat bin throughout the track but rather that there is a local discrepancy between the beat position in the flow layer and the beat layer of the composite auditory stream. The two drift apart for a moment, then come back together again. This is above all else a case of expressive timing, but it is such a marked rhythmic gesture that it might well have an impact on the entire structural framework of the flow. It indicates that the flow has been granted a significant level of rhythmic expressivity, perhaps to the extent of widening the thresholds for how far rhythmic structural categorical boundaries can be stretched, as was explained using Johansson’s (2010a) concept of ‘rhythmic tolerance’ in track B3 (page 189). Of course, this tolerance also depends on the immediately preceding rhythmic phrasing in the flow, and in this particular example, the passage leading up to the

---

175 See discussion in track A6 (page 120) on Danielsen’s concept of the ‘beat bin’ (2010b).
176 Depending on the musical context, these ‘strong positions’ need not be actual pulse-beat positions—they can be syncopated or placed on other off-beat positions which are construed as significant.
metrical anchor has introduced a significant amount of structural uncertainty, which makes the metrical anchor even more essential to defining the rhythmic structure.

Returning to the broader picture, it all comes down to an interplay between a rap track’s very formalised formal template and the flow’s variations within it. At different levels—from the large symmetrical blocks to the metrical identity and structure and on down to the microrhythm—the flow will reinforce or destabilise metrical and formal symmetry, present alternative counterrhythmic or crossrhythmic patterns and contribute to the tightness or looseness of the categorical rhythmic structure. The flow’s aesthetic identity derives from this constant play with both creating and challenging structural predictability. Clearly, there are many ways to accomplish this, leading us to the third sub-question, which we have already begun to answer: What are the essential rhythmic techniques—structural and expressive—which rappers employ in this interaction?

Adopting a term from Kjørup’s analysis of ‘verse-syntax counterpoint’ (2008), I described a variety of techniques which manipulate the interaction between linguistic structure (primarily syntax), rhyme position and musical metre as heteromorphies. These techniques utilise divergent metrical structure and tendencies towards divergence—using weak alternative lineations—and lend themselves to nearly limitless variation, but there are certain basic techniques to which rappers turn. Chief among them is the one-rhyme and its variations (presented in track B1), where a rhyme is placed on beat 1 rather than (or in addition to) the archetypical beat 4. This creates rhythmic tension by extending lines beyond their expected end points or tying a new line together with the preceding. The many varieties of one-rhyme techniques (like ‘added one-rhyme’, ‘half-line to one rhyme’, ‘repeated one-rhyming and so forth) have great nuance with regard to how severely and for how long they disrupt the expectation of convergent metrical structure, but their common denominator is that they revolve around beat 1 being a viable alternative to beat 4 for the placement of structurally significant (that is, ‘primary’) rhymes. There are other types of heteromorphies as well, like enjambement (well known from poetry) and early-onset rhyme, and various extended displacements of the line in relation to the musical metre (including extended half-line phrasing, extended one-rhyming and polyrhythmic or crossrhythmic phrasing). The use of divergent metrical structure and ambiguous lineation is an essential tool for rappers at the macrorhythmic level.

Another category of techniques which manipulate the boundaries of a ‘macrorhythmic’ structural parameter is the bridge rhyme and its variants. Rather than challenging the
symmetry of bars and groups of bars using line boundaries, bridge rhymes manipulate the boundaries of rhyme complexes and/or create connections between lines, bars or groups of bars through rhyme. In its most basic form, a bridge rhyme will bridge two separate structural units by extending the rhyme complex which coincides with one section into the next section, creating a different heteromorphy or ‘divergent structure’ where the boundaries of rhyme complexes and metrical structural units do not coincide. As shown in track B1, there are multiple variants of bridge rhymes and related techniques which destabilise and challenge structural symmetry to varying degrees. Common to all of them (like with the other heteromorphies) is the fact that they challenge structural predictability and symmetry. To use an intuitive analogy, if the flow is a river, techniques like bridge rhymes and heteromorphies are the variations in depth and width which make the current run at different speeds and take different paths as it makes its way from bend to bend.

These river bends which define the river’s overall form are analogous to the boundaries of the larger subsections of a rap verse. In track B2, I introduced several techniques which revolve around the gravity of the form—the flow, that is, will inevitably pass each major river bend, and how it does so is central to its expression and identity. One essential rap technique in this regard is the placement of grand gestures at certain focal points within a verse. For a gesture to be sufficiently ‘grand’, it must stand out from the musical context in some way. I described multiple variants in track B2. Some utilise variation in rhythmic phrasing, like crossrhythms and other stand-out patterns which stretch out across barlines or change density referent and subdivision type from quadruple to triple. Others use novel rhyme techniques, like weaker hidden rhymes, crossing rhyme classes or even non-rhyme—the latter triggering a ‘prediction error’ in the listener when the expected rhyme does not appear. Some grand gestures need not be features of ‘flow’ (in Edwards’s taxonomy of the term) but are rather techniques of delivery (like when rappers switch to a more speech-like tonality and/or rhythm in the final line of a verse) or even content (like a particularly dope rhyme or punchline).

Another essential technique related to formal gravity is cadencing—also presented in track B2—where rappers will build up rhythmic tension towards a resolution at the end of a verse or symmetrical sub-block. Typically, this is done by destabilising the symmetry to some degree first, before finally allowing the boundaries of lines, bars and rhyme complexes to coincide at the most prominent points of the formal structure. That a temporary or extended destabilisation of the symmetry within a verse gives the flow an accelerated, uninterrupted forward movement was evident in several examples in the first tracks of side B. The long
lines of André 3000 analysed in track B1 (from page 153) represent one example—the technique he employs is to repeatedly suggest potential line endings, then ignore them and continue the line further. The listener is primed for a break in the forward movement—a momentary stop where a new path forward is revealed. Instead, as the flow continues without stopping, the listener is dragged along on its unpredictable ride. Kendrick Lamar’s flow from ‘F*ck Your Ethnicity’, analysed in track B3 (from page 175), is a similarly wild ride—the formal symmetry is presented, then challenged, then dissolved. While the rapids are not dangerous and the current is not frantic, the flow takes a listener river rafting in the dark—one loses sight and track of the bends and is at the mercy of the flow until suddenly arriving at the river’s outlet—the end of the verse. Again, the techniques employed are based on manipulating the relationship between bars and lines and rhyme complexes, as well as variations in rhyme density. Also in play is a blurring of secondary and primary rhymes, wherein a listener is unsure of whether a new rhyme class is the start of a new, clearly defined structural section (a rhyme complex) or merely a detour within the one already underway. This type of technique is even more evident in the example from Nas’s ‘N.Y. State of Mind’ (track B1, from page 151). The muddying of boundaries between lines and rhyme complexes in relation to the musical metre is what propels this flow forward.

Up to this point, the ‘essential techniques’ I have presented are mostly of the structural kind. But the analyses on track B3 in particular highlight techniques which are primarily expressive. The overarching essential technique here is microrhythmic variation. The exact variation which a rapper’s microrhythmic phrasing imparts on a rhythmic event is so small that this is (in most cases) a technique based on ‘feel’ rather than a thought-out, deliberate, on-the-millisecond placement of syllables. Nevertheless, the intuitive application of such feel-based techniques is clearly very deliberate among many rappers, as was evident in Linni’s analogy of ‘waterskiing outside the wake’ (track B3, page 195), for example. When variation in microrhythmic phrasing is primarily expressive (a part of ‘delivery’), its structural contributions is towards a general ‘looseness’ of rhythmic categories at particular locations. It can vary in either direction (that is, being sometimes ahead of and sometimes behind the beat), as in the examples of ‘The Real Slim Shady’ (track A6, from page 125) and ‘Eple’ (track B3, from page 193). Microrhythmic variation can also be applied in a generally ‘consistent’ and systematic manner—for example, by being consistently ‘behind the beat’. This is a feature of Kendrick Lamar’s flow on ‘HiiiPower’, discussed in track B3 (from page 189), and it is also a stylistic feature of many rappers’ flow styles (Snoop Dogg comes to mind
as the archetypical example). These are examples of analogue features being (or becoming) *structural*.

In track B3, I introduced the concepts of *metrical anchors* and *bothness* as explanatory models for local structural implications of microrhythmic variations. These are not techniques in themselves, but they do shed light on some effects which a rapper might achieve when performing with noticeable timing variations. For example, a very ‘late’ placement of a metrical anchor, like the one in ‘Waterfalls’ in figure B3.1 (page 184), is a striking rhythmic gesture, particularly because the expressive timing in the passage preceding it makes the rhythmic structure unclear and ambiguous. This means that the metrical anchor appears to have been given carte blanche to define the position of the downbeat, even if it conflicts with what the musical background presents. Like the other essential techniques discussed above, microrhythmic phrasing—both expressive and structural—creates aesthetic rhythmic tension by blurring categorical boundaries and creating ambiguities. It can evidently also be used to make grand gestures.

The techniques described in track B4 are, like expressive microrhythmic phrasing, ways of using delivery to influence flow. In fact, of the melodic and rhythmic axes from speech-likeness to song-likeness, the latter axis is often characterised by a specific kind of expressive timing—a rhythmic phrasing which is more reminiscent of speech than of the more regular time-value classes of musical beats and their subdivisions. The techniques related to the melodic axis from speech to song, however, have other rhythmic-structural applications. For one thing, melodic patterns which are decidedly *not* speech-like, like some of the phrasings in Noname’s ‘Casket Pretty’ (from page 210) and the fully sung lines of Bizzy Bone from Bone Thugs-N-Harmony (figure B4.1, page 209), function like phonological and musical boundary markers. They serve as strong lineation evidence and demarcate structural units, even when these are otherwise destabilised by other rhythmic features. In the case of ‘Casket Pretty’, the lineation is ambiguous and divergent, and the melodic phrasing is an extra parameter which influences the intricate interplay of lines, bars and rhyme complexes and secondary rhyme connections. In a very different way, the melodic phrasing of Bizzy Bones creates a convergent metrical structure through melodic parallelism even though the syntax presents a divergent alternative. The final extended analysis of the thesis, looking at Chika’s ‘SONGS ABOUT YOU’ (from page 216), mainly focused on the use of contrasting sung and rapped

---

177 This is what Ohriner (2018, 2019c) calls ‘speech-rhythmic’ flow.
deliveries to create parallels between sections of different verses and demarcate structural blocks within verses. The analysis also expanded on the rap flow’s chimericity—that is, a rap flow typically consists of multiple vocal tracks blended together to a greater or lesser degree. In some flows, carefully balanced ‘backtracks’ or ‘dub tracks’ are nearly indistinguishable from the main vocal track—adding emphasis and power to specific words or syllables. At other times, the analogous chimera roars its heads, and two or more clearly distinguishable vocal streams move together as one. While Kendrick Lamar employs a gradual shifting of the volume balance between a clean main track and effects-processed dub tracks in ‘F*ck Your Ethnicity’, Chika uses a variety of techniques including straightforward dub track doublings at single words, a pitched-down ‘ogre voice’ doubling of lines at the ends of sections, and lines doubled by her four background singers in four-part harmony. This last type of five-headed chimera in particular shows how a rap flow is not just a single voice’s rhythmic output but an entire composite stream unto itself. The combination, blending and mixing of different vocal tracks comprise another essential technique in the rapper’s arsenal.

**Research question 2: How are rap flows analysed?**

The second main research question concerns the methodological, epistemological and even ethical dimensions of rap analysis. My answers to this question have been discussed and articulated at length throughout side A and exemplified on side B. I will now summarise what I believe are the most important takeaways from this thesis with regard to this question, which reads in full as follows: How are rap flows analysed? That is, how have the answers to the previous question [research question 1: ‘What are the core rhythmic features of rap flows?’] been produced, and how do these answers vary depending on the analytical framework in use?

In track A6, I stated that I believe the approach which has arisen in the field of rap analysis at large—developing ‘ad hoc’ methods for representing the specific features which are relevant to the analysis at hand—is a richer, more fruitful approach than looking for ‘universal methods’ which can be standardised for all types of rap flow analysis. This is not to say that one should ignore the innovations of previous scholarship, or that representational originality is a goal in itself. My point is rather that different tools are suitable for different tasks—and that this is A Good Thing. The methodological meta-framework I suggest in this thesis is derived from the correspondences between different ‘levels’ of rhythm perception and different methods for the visualisation and representation of features of rap flows. These
different zoom levels can be applied both separately and in tandem depending on the analytical task. From the widest to the narrowest lens, these are as follows:

1) **The global zoom level:** concerns itself with ‘global features’ of rap flows, where multiple flows and/or parts/features of flows are superimposed, added, averaged and/or compared to one another without any individual piece of flow being visualised. Examples of this level of representation and analysis are the graphs and tables of the various corpus studies presented in track A2 (as well as the ones used to analyse rhyme presented in track A5).

2) **The macrorhythmic zoom level:** eschews information about subdivisions of the beat completely, and even the beat level might only be represented in an approximate manner, depending on how relevant it is to the specific analysis. Lyrical transcription divided into lines is perhaps the ‘sparsest’ type of macrorhythmic representation and omits representation of the beat positions, while different types of ‘flow diagrams’ might show beat positions to different degrees of exactness.

3) **The quantised zoom level:** shows quantised rhythmic units down to the smallest units of subdivision, but it does not visualise any within-category variation—that is, microrhythmic features and/or expressive timing. Traditional music notation and the various grid-based representations presented in track A2 are common systems for representation at this zoom level.

4) **The microrhythmic zoom level:** as the level of greatest zoom, it aims at representing the rhythmic features which happen within (and between) the smallest quantised rhythmic categories (that is, the subdivisions of the beat). Microrhythmic features—whether they are structural (for example, consistent ‘behindness’) or expressive—are often central to a rap flow’s aesthetic, and these are represented visually either in relation to some sort of ‘ruler’ (either an isochronous or idealised temporal grid representing canonical rhythmic categories or other aspects of the rhythmic context such as parts of the musical background) or the internal relationship between the rhythmic events.

In addition to these categories of rhythmic representation, I both address and apply certain types of visualisation which represent melodic content. These can either be quantised, as in traditional notation which shows discrete (that is, digital) musical pitches, or continuous (analogue), as in the pitch contours used in linguistics.
Because one of the goals of this thesis was to evaluate and contribute to methodological approaches highlighting features which are ‘under-analysed’ in the field, I decided to focus mainly on the macrorhythmic and microrhythmic zoom levels.

Turning first to the macrorhythmic zoom level, the most important macrorhythmic parameters to represent and analyse are formal structure, the relationship between lines and bars—the ‘metrical structure’, which might be convergent (lines and bars coincide) or divergent (they do not)—and the structure of rhymes and various rhyme connections. It is therefore crucial that macrorhythmic representation visualise the interactions among lines, bars, primary and secondary rhymes and subsections of a verse. Ideally, this should be done in a non-hierarchical manner.

Traditional notation will divide lines according to musical metre, and this might obfuscate the line as a significant structural parameter. In general, in fact, the inherent prioritisation of musical metre in traditional notation is an issue. For representations of full verses or longer segments—for example, figures B1.12 (page 151) and B2.5 (page 171)—I have therefore chosen a type of representation which flips this hierarchy. The dominant lineation of the lyrical lines is transcribed in accordance with my interpretation and/or how the rappers themselves transcribe their lines (discussed in track A4, from page 67), and these lines are then reorganised and superimposed onto the musical metrical framework. Of course, this is also a hierarchisation of the two interacting constituents of lines and bars, but this time in favour of the line over the musical metre. My reasoning for this shift is that the musical metre is the stable, unchanging part of the joint metrical structure, and the most significant feature to represent visually is the line. The flow’s identity and expression are defined through the lineation and how it interacts with the more predictable musical metrical background, and it is more representative of the listening experience to show how the structuring power of the regularity of the musical metre is challenged by the lineation. Visualisations prioritising musical metre diminish the structuring power of lineation, and I believe prioritising lines is a more faithful rendition of the listening experience.

Of course, a continuous graphic line which prioritises neither would be the ultimate means of emphasising a non-hierarchical relationship, but that is not feasible in print. Some variations of circular plotting (like Page’s ‘flowprints’, discussed in track A2) exist, but these are very difficult to parse, and visual clarity is paramount. Video (that is, a moving picture) is a fantastic alternative, particularly as it can be combined with audio, and the application of audiovisual representation to music is a solution to many of the problems inherent in graphic
representation. My solutions for graphic representation have all attempted to avoid line breaks where possible or used line breaks at the places where the boundaries of lines and bars align.

At the microrhythmic zoom level, the key epistemological point to the representations is that they communicate both microrhythm and the quantised categorical rhythmic structure which the microrhythm creates, challenges and coexists with. Whether the microrhythmic phrasing is consistent and structural, ‘consistently inconsistent’ and expressive, or expressive in a local sense, its effects and identity are tied to the categorical structure. The specific microrhythmic phenomena and techniques which I have analysed are characterised by some sort of categorical ambiguity or ‘categorical friction’ between layers of the composite auditory stream, which means that a key feature of the representation is the various alternative categorical interpretations of same.

Microrhythmic analysis is also highly dependent on the accompanying discussion and description of the rhythmic experience, and many of the most significant analytical points arise in the nuances of different valid structural interpretations and emergent reinterpretations. My representations do not differ significantly in content from those of other scholars (for example, Ohriner’s, which also include potential quantised interpretations), but my analytical points are perhaps more phenomenologically oriented. The focus is less on the exact positions of rhythmic events than on the ways in which they can illuminate our rhythmic experience—and in particular how it is ambiguous, changing, emergent and individual. If one were to take away one single insight from my microrhythmic analyses, it should be that one must embrace and emphasise the inherent inexactness and ambiguity of the relationship between the acoustic signal and the perceived rhythmic structure.

**Rap analysis: What does it do, and what can it do?**

This leads me to my final point. What are the potential applications of the type of rap analysis performed in this thesis with regard to, respectively, future scholarly research of the inter- or cross-disciplinary type, practical music making, music education, and the informal dissemination and ‘criticism’ performed by the particularly interested hip-hop head? While this analytical work exists within and derives from a very specific academic tradition, it is influenced to a great degree by the other cultural and professional arenas to which I belong, and I both hope and believe that some of the knowledge and methodology produced in this thesis can and will be useful outside of the musicological field.
First, I have made some forays into the interdisciplinary in this thesis. Methods and theory from the fields of literature and linguistics have been essential to forming my analytical approach. The relationships among language and poetry and music are at the core of the artistic expression of rap, both in general and for rap *flows* in particular. Further explorations of cross-disciplinary approaches would be a natural continuation of the work in this thesis. The relationships among poetic lines and musical metre, convergent and divergent metrical structure, rhyme complexes, primary and secondary rhymes, and a (even more) nuanced analysis of *lineation* in rap are natural meeting points for all three disciplines. Similarly, the analysis of the *topography of rhythmic events*, as discussed in track A6 (from page 109), with several levels of accents/stress—using methods from literary and linguistic analysis—could potentially further our understanding of the intricacies of the rhythmic structure (at the quantised *zoom* level) of rap flows. The perhaps most obvious avenue for developing interdisciplinary approaches to the analysis of rap, however, is the analysis of rap’s melodies. The groundwork laid by scholars like Gilbers et al. (2019), Ohriner (2019a) and Hognestad (2022) will support important future research into rap as a type of singing with different speech-like characteristics. While rap adheres to discrete (quantised, digital) musical structure both melodically and rhythmically, there is a great deal of nuance to unpack in how different types of ‘speech-likeness’ work with and against the norms of musical structure. My exploration of the topic in track B4 introduces some ideas and some potential paths forward.

With regard to the more practical and educational application of the concepts, techniques and methods presented in this thesis, there are some obvious uses, some less obvious ones, and some potential ethical dilemmas which merit a short discussion. From a practitioner’s standpoint, the identification, naming and nuancing of specific techniques are clearly useful. Most rappers are intuitively aware of and proficient in using techniques like *one-rhyming*, and introducing a term for it is simply a way to ease communication. Reflecting on topics like lineation, expressive timing, speech-likeness and chimericity might also be useful for rappers’ creative endeavours. Certain techniques and concepts might also inform the work of producers, mixing engineers and/or ‘beatmakers’—all of whom are more or less familiar with and take account of, for example, the gravity of the form, cadencing, microrhythm and various ways of using dub tracks and chimericity. Slightly more problematic are the questions concerning representational tools and any type of formalisation of the communication of certain concepts in educational contexts. For one thing, there is the question of whether one should ‘teach rap’ in a formalised, organised music education sense in the first place. The hip-
hop ethos of ‘each one teach one’, of being in and communicating a culture and its values, as well as a more holistic view of the culture’s constituents, stands in contrast with the traditional ‘music school’ type of formalised communication. And even if one believes it is appropriate to, for example, create rap courses or textbooks, the ethical considerations of which content and which types of language and systems for visual representation should be applied are fraught.

I will not put my hand into that hornets’ nest now—there is no way to treat the topic with the rigour and thoroughness it deserves in only a few paragraphs—but I will comment on some of the implications which might arise from the application of formalised representational systems in more practical contexts, as I have no doubt that this will (and does) happen, whether it should or not. Fundamentally, all communication and description of musical features which are abstracted to any extent from the music itself will emphasise some features or elements of the musical whole over others. This has been one of the main criticisms of using traditional notation to represent rap—there are so many significant features of rap’s rhythmic structure which are either ignored or obfuscated to some degree in simple quantised representation. Of course, I have proposed solutions to this with the zoom levels and the different types of representation I apply, but there is no doubt that as rap finds its way into contexts where the goal is not to give the most accurate and culturally sensitive rendition of a piece of recorded music, problems might arise. There is a significant difference between the use of quantised representation as a descriptive tool—highlighting specific musical features—and its common application as a prescriptive tool to communicate how music should/will be performed. The latter is a challenge for music educators, arrangers and composers (and even potentially some rappers) as it becomes more and more common in the context of ‘rap scores’—for example, in musicals (like Lin-Manuel Miranda’s Hamilton) or in the countless books with lead sheet versions of mainstream pop hit songs (which now often feature a rapped verse). Does it make sense to use prescriptive notation to communicate rhythmic structures which feature significant categorical ambiguity due to microrhythmic phrasing/expressive timing, for example, and would the notation be different in descriptive contexts? There are examples of transcriptions of rap flows which attempt to make a MIDI playback of the notation fit as closely as possible with the P-centres of the flow’s syllables (like some of Connor’s work, as described in track A2). In passages with significant microrhythmic variation, this approach leads to nonsensical notation which completely foregoes any information about rhythmic categories in favour of treating bars and beats as absolute
durational categories instead. There are some proposed solutions for prescriptive notation of ‘expressively timed’ music, like Fernando Benadon’s ‘gridless beats’ (2009), but it seems improbable that any such system will be adopted by the commercially driven publishing of scores for the wider public. Should one perhaps just accept that microrhythmic features and techniques cannot be formalised as rhythmic structure on paper or screen, and that their production is dependent on performers’ expressive interpretation or a reproduction of recorded examples? This would in turn mean that the communication of essential microrhythmic features and details in the music is entirely dependent upon how conscientious the performer (or the teacher using the score in an educational context) is about internalising and reflecting upon these aspects of rap performance.

Of course, I can do little about how the communication and representation of rap flows will develop, but if my work can have any sort of positive contribution to anyone’s understanding or enjoyment of this magical rhythmic expression, I will be satisfied. Whether that person is a music scholar, a rapper, a producer, a music educator, a hip-hop head or someone like myself—who is all of the above—I hope that they feel like they are a bit closer to knowing—with confidence and hip-hop swagger—what makes the shit dope.
List of figures

p. 4 Figure A0.1: ‘What makes the shit dope?’ Hand-drawn mega-question.

p. 64 Figure A4.1: Correspondence of musical and poetic metrical-structural categories.

p. 65 Figure A4.2: Transcription of bars 1–4 of J Cole’s verse on 21 Savage’s ‘a lot’.

p. 71 Figure A4.3: Geto Boys: ‘Mind Playing Tricks on Me’. Flow diagram of first 4 bars of verse 3 (by Scarface). From Gudnason (2016).


p. 72 Figure A4.5: Picture of two pages from one of the author’s notebooks.

p. 77 Figure A4.6: Different music-rhythmic realisations of repeated dactylic feet.

p. 80 Figure A4.7: Divergent metrical structure with consistent symmetrical rhythmic groupings. Dizzee Rascal: ‘Dirtee Cash’.

p. 97 Figure A6.1: Periodicities in Warren G and Nate Dogg’s ‘Regulate’.

p. 104 Figure A6.2: Transcription of bars 1–2 of J Cole’s verse on 21 Savage’s ‘a lot’.

p. 110 Figure A6.3: Linguistic prosody. Hierarchy of lexical stress.

p. 113 Figure A6.4: A topography of rhythmic events.

p. 123 Figure A6.5: Macrorhythmic representation. Eminem: ‘The Real Slim Shady’.

p. 124 Figure A6.6: Quantised representation (musical notation). Eminem: ‘The Real Slim Shady’.

p. 126 Figure A6.7: Microrhythmic representation (waveform, spectrogram, lyrics and grid). Eminem: ‘The Real Slim Shady’.

p. 127 Figure A6.8: ‘Microrhythmic rhyme’ in Eminem’s ‘The Real Slim Shady’.

p. 128 Figure A6.9: Examples of ‘even’ sixteenth note patterns in the flow in Eminem’s ‘The Real Slim Shady’.

p. 129 Figure A6.10: Screengrab from Melodyne (with superimposed lyrics) showing pitch contour of two phrases from Eminem’s ‘The Real Slim Shady’.

p. 137 Figure B1.1: Enjambment in Snoop Dogg’s ‘Gin And Juice’.

p. 137 Figure B1.2: Simple one-rhyme. Nino Durden on Jon Krem’s ‘Fucker Med Krem (G-Mix)’.

p. 139 Figure B1.3: Half-line rhyme to one-rhyme in East 17’s ‘Deep’.

p. 139 Figure B1.4: One-rhyme in dual-planed trap metre. Bad Bunny on Cardi B’s ‘I Like It’.

p. 141 Figure B1.5: Added one-rhyme (and split rhyme) in Diaz’ ‘Crossing Borders’ (with Petter).

p. 142 Figure B1.6: One-rhyme variations in Ms. Lauryn Hill’s ‘Doo Wop (That Thing)’.
p. 144 Figure B1.7: Different hypothetical placements of simple one-rhymes within hypermetre.

p. 146 Figure B1.8: Bridge rhyme using simple added one-rhyme. Big Boi on OutKast’s ‘Ms. Jackson’.

p. 147 Figure B1.9: Bridge rhyme in South Park Mexican’s ‘High so High’.

p. 148 Figure B1.10: Pivot rhyme in Side Brok’s ‘Setra’.

p. 149 Figure B1.11: Early primary rhyme onset and bridge rhyme. André 3000 on OutKast’s ‘Skew It on the Bar-B’.

p. 151 Figure B1.12: Various bridge rhyme techniques in Nas’s ‘N. Y. State of Mind’.

p. 154 Figure B1.13: Successive reinterpretation of dominant lineation. André 3000 on OutKast’s ‘Rosa Parks’.

p. 156 Figure B1.14: Metrical structure of excerpt of André 3000’s first verse (the track’s second) on OutKast’s ‘Aquemini’.

p. 159 Figure B1.15: Ambiguous lineation evidence and successive bridge rhymes. André 3000 on OutKast’s ‘Aquemini’.

p. 160 Figure B1.16: Primary and secondary rhymes. André 3000 on OutKast’s ‘Aquemini’.

p. 163 Figure B2.1: Grand opening gesture of Gatas Parliament’s ‘Stem Gatas Parlament’.

p. 165 Figure B2.2: Grand opening gesture of Big Boi’s verse on OutKast’s ‘Skew It on the Bar-B’.

p. 168 Figure B2.3: Divergent metrical structure and cadencing. MF DOOM: ‘THAT’S THAT’.

p. 169 Figure B2.4: Cadencing through increased rhyme density. Side Brok: ‘Heimegut’.

p. 171 Figure B2.5: Metrical structure, beat/rhyme positions, hypermetric structure and novel techniques in verse 1 of Lars Vaular’s ‘Helt om natten, helt om dagen’.

p. 172 Figure B2.6: Musical notation transcription with marked rhymes of bars 5–8 of Lars Vaular’s ‘Helt om natten, helt om dagen’ from Oddekalv (2017).

p. 174 Figure B2.7: Hidden crossing rhymes. Lars Vaular: ‘Helt om natten, helt om dagen’.

p. 178 Figure B2.8: Metrical, hypermetric and rhyme scheme structure of the first verse of Kendrick Lamar’s ‘F*ck Your Ethnicity’.

p. 183 Figure B3.1: Metrical anchor in TLC’s ‘Waterfalls’.

p. 186 Figure B3.2: Metrical anchors and rhythmic tolerance in Side Brok’s ‘Si Eiga Rås’.

p. 189 Figure B3.3: Metrical anchors creating categorical ambiguity in Kendrick Lamar’s ‘HiiiPower’

p. 193 Figure B3.4: Bothness in Neste Planet’s ‘Eple’.

p. 196 Figure B3.5: Instrumental track, vocal track, spectrogram with superimposed lyrics and metrical anchors in Chance The Rapper’s ‘Do You Remember’.
p. 208 Figure B4.1: Melismata in Bizzy Bones’ verse on Bone Thugs-N-Harmony’s ‘Tha Crossroads’. Transcription using music notation.

p. 210 Figure B4.2: Microrhythmic analysis of Noname’s ‘Casket Pretty’, backing track and vocal track waveforms and vocal track spectrogram with overlaid position of metrical anchors.

p. 211 Figure B4.3: Various structuring parameters in the first part of the verse of Noname’s ‘Casket Pretty’.

p. 212 Figure B4.4: Nested duple cross-rhythm in Noname’s ‘Casket Pretty’.

p. 213 Figure B4.5: Screengrab from Melodyne showing a phrase from Noname’s ‘Casket Pretty’.

p. 214 Figure B4.6: Screengrab from Melodyne showing another phrase from Noname’s ‘Casket Pretty’.

p. 216 Figure B4.7: Transcription of the first six bars of the first verse in Chika’s ‘SONGS ABOUT YOU’ using music notation.

p. 217 Figure B4.8: Song structure of Chika’s ‘SONGS ABOUT YOU’.
Playlist

21 Savage (2018), a lot. On \textit{i am > i was}. Slaughter Gang LLC/Epic Records


Cardi B (2018), Bodak Yellow. On \textit{Invasion of Privacy}. Atlantic Records

Cardi B (2018), I Like It. On \textit{Invasion of Privacy}. Atlantic Records

Chance the Rapper (2019), Do You Remember. On \textit{The Big Day}. Chance the Rapper LLC

Chika (2020), SONGS ABOUT YOU. On SONGS ABOUT YOU. CHIKA / Warner


D’Angelo (2014), 1000 Deaths. On \textit{Black Messiah}. RCA Records


Dizzee Rascal (2011), Dirtee Cash. On Tongue N’ Cheek. Dirtee Stank Recordings


Emile The Duke (2015), Fredag. On “\textit{Aasen}”. Mook Time


Gatas Parlament (2002), Ti Mot En. On \textit{Holdning over Underholding}. Tee Productions


Kendrick Lamar (2011), F*ck Your Ethnicity. On \textit{Section.80}. Top Dawg Entertainment / Section.80

Kendrick Lamar (2011), HiiiPower. On \textit{Section.80}. Top Dawg Entertainment / Section.80


Lars Vauler (2010). Helt om natten, helt om dagen. On Helt om natten, helt om dagen. NMG/G-Huset

MF Doom (2009). THAT’S THAT. On BORN LIKE THIS. Lex Records Ltd


The Streets (2002). The Irony of It All. On Original Pirate Material. Pure Groove Ltd trading as Locked On

TLC (1994). Waterfalls. On Crazysexycool. LaFace Records LLC

Bibliography


Horn, E. A. (2010). *Poetic organization and poetic license in the lyrics of Hank Williams, Sr. and Snoop Dogg*. PhD dissertation, University of Texas at Austin.


