

Muzikaal Verhaal

Een onderzoek naar de mogelijkheden van het karakteriseren
van muziek in termen van verhaalstructuur

Maartje Masereeuw
Studentnummer: 4172558
m.masereeuw@students.uu.nl

Bachelorscriptie 7,5 ECTS
Kunstmatige Intelligentie
Universiteit Utrecht

Begeleiders:
Rick Nouwen
Cynthia Liem

17 augustus 2016

Inhoudsopgave

Inleiding	2
1 Gerelateerd werk	4
1.1 Music Information Retrieval: bestaande technieken	4
1.1.1 Music Information Retrieval	4
1.1.2 High-level versus low-level	4
1.2 Muziek en cinematische scènebeschrijvingen	5
1.3 Mogelijkheden van tekstuele analyse	5
1.3.1 Gemeenschappelijke narratieve elementen	5
1.3.2 Vendler's werkwoordclassificatie	6
1.3.3 Doelgerichtheid en events	7
1.3.4 Toepassing door Liem et al. (2013)	8
1.4 Verbinding aan de muziek	9
1.4.1 Eigen redenen van deelnemers	9
1.4.2 Muzikale verwachting	10
1.4.3 Voorspelbaarheid	11
1.4.4 Verrassing	12
1.4.5 Spanning	13
1.4.6 Samenvattend schema muzikale verwachting	13
2 Methode	15
2.1 Algemeen	15
2.2 Onderzoeksgroep	15
2.3 Meetinstrumenten	16
2.3.1 Classificaties en fragmenten	16
2.3.2 Audio-analyse	17
2.4 Procedure	17
2.4.1 Stap 1: onderliggende verhaalstructuur	17
2.4.2 Stap 2: koppeling tekst en muziek	20
3 Analyse	21
3.1 Tekstanalyse	21
3.2 High-level audio-analyse	24
3.3 Low-level audio-analyse	26
3.4 Koppeling classificatie - Huron (2007)	29
4 Resultaten	31
5 Discussie en suggesties voor vervolgonderzoek	32
5.1 Discussie	32
5.1.1 Classificatie	32
5.1.2 Chronologie en klasse 1	32
5.1.3 Segmentatie	33
5.1.4 Verbinding tekst en muziek	33
5.2 Suggesties voor vervolgonderzoek	33
5.2.1 Vraagstelling	33
5.2.2 Tekstanalyse	34
5.2.3 Vertaling naar muzikale verwachting	34
5.2.4 Low-level vervolg	34
6 Samenvattende conclusie	35
Referenties	37

Inleiding

Muziek is overal, in de meest uiteenlopende contexten. Muziek vormt een wisselwerking met en dient als ondersteuning bij allerlei disciplines, kunstvormen en situaties. Doordat het zo breed ingezet wordt en alomtegenwoordig is, krijgt muziek al snel bepaalde associaties met zich mee. Dit gegeven wordt uitgebreid uitgebuit in multimediaproducties waarin beeld, geluid en tekst geïntegreerd zijn, zoals film, theater en televisie.

Ook amateurs thuis kunnen hierin geïnteresseerd zijn en muziek op dezelfde manier willen gebruiken. Denk bijvoorbeeld aan iemand die een filmpje heeft gemaakt van zijn kat en dat op YouTube wil zetten met een bepaalde achtergrondmuziek om de juiste sfeer te creëren. Hoe vindt hij die muziek, zonder een uitgebreid vocabulaire aan muziektheoretische termen?

Tegenwoordig, in het tijdperk van digitalisering, is een gigantisch aandeel van alle bestaande muziek online beschikbaar. Het overzichtelijk houden van al deze data eist veel van zoeksystemen. De persoon van het kattenfilmpje zou uitgebreid op Spotify allerlei muziek kunnen luisteren tot hij precies de juiste muziek voor zijn kattenavontuur heeft gevonden. Zou het echter niet makkelijker zijn om in muziekdatabases te kunnen zoeken op basis van de verhaalstructuur van het filmpje?

Deze scriptie vloeit voort uit deze vraag. Ik onderzoek het verband tussen muziek en de verhaalstructuur van cinematische scènebeschrijvingen. Als dit verband sterk genoeg is, is het in de toekomst wellicht mogelijk om muziek te karakteriseren in termen van een narratieve multimediacontext. Zo zouden misschien ooit kattenliefhebbers simpelweg kunnen zoeken naar muziek bij hun filmpje door het avontuur van hun kat in woorden te beschrijven.

Van veel systemen die informatie uit muziek halen is de werking gebaseerd op methoden uit de Kunstmatige Intelligentie. Om informatie uit muziek te halen (zowel uit audio- als visuele bronnen) wordt veel gebruik gemaakt van statistische modellen. Bovendien komt er veel psychologie kijken bij theorieën over verwachtingspatronen en beweging in muziek. Een taalkundige aanpak is relatief nieuw in het veld en wellicht veelbelovend.

Liem et al. (2013) concluderen uit hun onderzoek dat het connotatieve verband tussen spontane, vrije-vorm beschrijvingen van verhalende beelden en muzikale informatie inderdaad sterk genoeg is om te herkennen en te generaliseren. In deze scriptie onderzoek ik hoe ze dat hebben kunnen concluderen en verdiep ik me in wat het verband precies is. Dit doe ik aan de hand van de volgende vragen:

Deelvragen theoretisch kader:

1. Wat is de positie van dit onderzoek binnen het onderzoeksveld?
2. Wat zijn de mogelijkheden voor een tekstuele analyse van vrije-vorm beschrijvingen van verhalende beelden?
3. Wat zegt literatuur over muzikale verwachting?

Deelvragen praktijkonderzoek:

4. Is er een “gemiddelde” onderliggende verhaalstructuur in de cinematische beschrijvingen bij een muziekstuk?
5. Op welke manier kan de tekstuele verhaalstructuur verbonden worden aan specifieke elementen in de muziek?

Hoofdvraag:

6. Wat zijn veelbelovende resultaten uit mijn onderzoek, die meegenomen kunnen worden in vervolgonderzoek?

De eerste drie deelvragen behandel ik in hoofdstuk 1. In paragraaf 1.1 plaats ik dit onderzoek in het brede onderzoeksveld van Music Information Retrieval. In paragraaf 1.2 beschrijf ik het onderzoek van Liem et al. (2013) waar deze scriptie op voortbouwt, om in paragraaf 1.3 de werkwoordclassificatie van Vendler (1967), het fenomeen *events* en de classificatie van Liem et al. (2013) te behandelen. In paragraaf 1.4 beschrijf ik hoe Liem et al. (2013) hebben geprobeerd tekst en muziek te verbinden en haal ik de muzikale verwachtingstheorie van Huron (2007) erbij.

In hoofdstuk 2 beschrijf ik de methode van mijn praktijkonderzoek en zoek ik antwoord op deelvraag 4 en 5. In paragraaf 3.1 pas ik een tekstuele analyse toe op een aantal teksten die cinematische scènes beschrijven waar bepaalde muziek onder zou kunnen passen. Hiervoor gebruik ik mijn eigen aangepaste versie van de classificaties van Vendler (1967) en Liem et al. (2013). Vervolgens analyseer ik in paragraaf 3.2 en 3.3 de audio die bij die teksten hoort.

Vervolgens zet ik in hoofdstuk 4 belangrijkste resultaten van mijn onderzoek uiteen, zodat ik in hoofdstuk 5 mijn hoofdvraag kan beantwoorden. Dat doe ik door middel van een kritische discussie van mijn eigen onderzoek en het geven van suggesties voor vervolgonderzoek. Ik eindig mijn scriptie met een samenvattende conclusie.

1 Gerelateerd werk

1.1 Music Information Retrieval: bestaande technieken

1.1.1 Music Information Retrieval

Dit onderzoek bevindt zich in het zeer interdisciplinaire onderzoeksveld van *Music Information Retrieval* (MIR). Die term staat simpelweg voor alle strategieën die proberen informatie te verkrijgen uit muziek - zowel uit audiosignalen als uit symbolische bronnen (bladmuziek). Het doel is om muziek, of informatie over muziek, makkelijker te vinden te maken. Dit is een groot en groeiend vakgebied met veel betrokken partijen, zoals academische en commerciële laboratoria, archieven en bibliotheken. Er is een snelle ontwikkeling gaande, om de verwachtingen bij te kunnen houden op het gebied van zoek- en browsefunctionaliteit.

Casey et al. (2008) geven in hun artikel een uitgebreid overzicht van de huidige technieken en uitdagingen in het werkveld. Ze besteden vooral aandacht aan de methoden waarbij informatie over audiosignalen en muzieksymbolen automatisch verkregen en verwerkt kan worden voor het gebruik in zoeksystemen. Ze maken onderscheid tussen verschillende methoden, namelijk methoden gebaseerd op metadata en methoden gebaseerd op muzikale inhoud, waarbij bovendien onderscheid wordt gemaakt tussen high-level en low-level kenmerken.

De meest gebruikte methode om toegang te krijgen tot muziek werkt op basis van tekstuele metadata. Dat is geen informatie die te vinden is in de audio of symbolische weergave van de muziek zelf, maar informatie die door derden is opgeschreven. Casey et al. (2008) maken onderscheid tussen feitelijke en culturele metadata. Feitelijke metadata bevat objectieve waarheden over een track; denk bijvoorbeeld aan titel, componist of jaartal. Culturele metadata bevat subjectieve concepten, zoals bijvoorbeeld genre, klankkleur of de emotie die het uitdraagt. Omdat culturele metadata subjectief is, brengen die methodes nog wel eens problemen met zich mee; hoe meer data, hoe groter de kans op inconsistente beschrijvingen of variatie in de coderingen. Een goede redactionele procedure is dan nodig; eventueel met behulp van sociale media.

1.1.2 High-level versus low-level

Naast de op metadata gebaseerde methoden zijn er methoden gebaseerd op de muzikale inhoud zelf, zoals informatie uit het audiosignaal of symbolische bronnen (bijvoorbeeld bladmuziek). Op inhoud gebaseerde methoden zijn volgens Casey et al. (2008) geen vervanging, maar verbeteringen van de metadatamethoden en kunnen identificeren wat de gebruiker zoekt, zelfs als die niet precies weet waarnaar hij op zoek is. Veel van deze methoden gebruiken technieken uit de Kunstmatige Intelligentie (zoals machine learning software).

Casey et al. (2008) maken onderscheid tussen high-level en low-level kenmerken. High-level muzikale kenmerken zijn elementen van muziek die te vangen zijn in muzikale concepten als melodie en harmonie. Het gaat hier om het soort kennis dat een gevorderde luisteraar zou hebben over een muziekstuk, gebruik makend van traditionele (Westerse) muziekconcepten. Low-level kenmerken bestaan uit informatie die rechtstreeks te halen is uit een digitaal audiosignaal, zonder dat daar noodzakelijk menselijke kennisconcepten aan te pas komen.

De moeilijkheid van Music Information Retrieval zit in het overbruggen van de kloof tussen low-level en high-level kenmerken, tussen audiosignaal en muziektheoretische concepten. De bedoeling is dat de muziek op zo'n manier geëncodeerd wordt dat de schema's dat de schema's overeenkomen met muzikale domeinkennis, zodat ze vervolgens gebruikt kunnen worden in zoeksystemen.

Het onderzoek waar ik deze scriptie op baseer maakt een verbinding tussen muziek en verhaalstructuur. Ik analyseer metadata en beschrijf de high-level kenmerken van de bijbehorende muziek. Dat kan in de toekomst het identificeren van de juiste low-level kenmerken gemakkelijker maken.

1.2 Muziek en cinematische scènebeschrijvingen

Dit onderzoek bouwt voort op het onderzoek van Liem, Larson & Hanjalic (2013). Zij hadden een database aan productiemuziek, vrijgegeven zonder royalty's. Ze vroegen de deelnemers aan hun onderzoek om bij een fragment een cinematische scène te beschrijven waar de muziek goed onder zou passen. De deelnemers kregen geen vaste regels of vocabulaire die ze moesten gebruiken voor hun beschrijving, maar ze hadden alle vrijheid om hun eigen woorden te gebruiken. Pas achteraf werd het taalgebruik van de verkregen tekstjes geanalyseerd.

Om te testen of de beschrijvingen überhaupt aan de muziek te verbinden zijn, hebben ze een tweede experiment gedaan: ze lieten andere deelnemers een tekstje lezen en gaven hen een aantal fragmenten. De deelnemers moesten vervolgens het fragment kiezen dat ze het beste bij de tekst vonden passen, om te testen of dat inderdaad het fragment was waarbij het tekstje geschreven was. Dit werd zeer goed gedaan. Dat betekent dat het verband tussen tekst en muziek in ieder geval door menselijke oren te leggen is.

Vervolgens hebben Liem et al. (2013) de teksten geanalyseerd, aan de hand van drie vragen:

1. Zijn er gemeenschappelijke tekstuele elementen in de verschillende teksten, per individueel fragment?
2. Zijn er in het algemeen gemeenschappelijke elementen zichtbaar in alle teksten?
3. Kunnen gebruikers zelf aangeven waarom ze een bepaalde scène beschreven hebben?

De eerste twee vragen behandel ik in de volgende paragraaf (1.3). Liem et al. (2013) hebben er onder andere de taalkundige eventtheorie van Vendler (1967) bijgehaald. Ik bespreek hoe ze dat gedaan hebben en wat mogelijke alternatieven kunnen zijn. De derde vraag gaat over de redenen voor de keuze voor bepaalde beelden bij de muziek. Daar ligt mogelijk iets interessants voor de link tussen muziek en verhaalstructuur. In paragraaf 1.4 bespreek ik de uitkomsten van Liem et al. (2013) en haal ik de ideeën van Huron (2007) erbij over verwachting in muziek. Zo kan ik de intuïties van de deelnemers ondersteunen met muziekpsychologische theorie.

1.3 Mogelijkheden van tekstuele analyse

1.3.1 Gemeenschappelijke narratieve elementen

Om gemeenschappelijke elementen te vinden in de vocabulaire van een groot aantal tekstjes per individueel fragment, hebben Liem et al. (2013) zogenaamde “word clouds” opgesteld. Daarmee wordt weergegeven welke woorden in een tekst voorkomen; woorden die veel voorkomen worden groter en duidelijker weergegeven. Ze hebben hiervoor eerst veelgebruikte stopwoorden en het woord “scène” verwijderd.

Zo'n word cloud kan inzicht geven in het taalgebruik van luisteraars. Zo geeft de word cloud van het fragment *Exciting Trailer* (componist Kevin MacLeod) weer dat onder andere de woorden “battle”, “people”, “army” en “preparing” vaak voorkomen. Veel van de luisteraars krijgen een gevoel van “voorbereiding voor een machtssituatie”. Als het gaat over hoofdpersonen, komen de woorden “fighters” en

“heroes” naar voren. Wat betreft locaties worden “Europa” en “USA” het meest genoemd.

Liem et al. (2013) probeerden ook breder te generaliseren, door de kijken naar overkoepelende elementen bij álle teksten van álle fragmenten. Ze gebruikten als definitie van narratief “een gesproken of geschreven uiteenzetting van gerelateerde gebeurtenissen, een verhaal”. Dit zagen zij als motivatie om gebeurtenissen, of *events*, als bouwstenen te zien van een beschrijving. Ze hebben daarom inspiratie geput uit de eventtheorie, en specifiek het werk van Vendler (1967).

Ze beschrijven de vier klassen van werkwoorden die Vendler (1967) onderscheidt op basis van het verschil in hun temporele eigenschappen. Liem et al. (2013, p. 26) beschrijven die klassen als volgt:

Tabel 1: Vendler volgens Liem et al. (2013)

State	Geen interne structuur
Activity	Wel interne verandering maar geen eindpunt
Achievement	Een plotseling eindpunt
Accomplishment	Een periode van opbouw en vervolgens een eindpunt

Gebruik makend van deze definities van klassen, viel het op dat er bij veel fragmenten een algemene overeenstemming leek te bestaan over in welke klasse de muziek viel. Ik zal nu ingaan op hoe Vendler (1967) zelf zijn vier klassen omschrijft, de verschillen met de versie van Liem et al. (2013) en de mogelijke alternatieven die bruikbaar zijn voor dit onderzoek.

1.3.2 Vendler’s werkwoordclassificatie

Muziek bestaat niet zonder het concept tijd. Een enkel moment van een muziekfragment is slechts een geluid; het wordt pas muziek als de tijd verstrijkt en er op een bepaalde manier een ontwikkeling gaande is. Na het luisteren is het vervlogen. Ook bij taal is het concept tijd belangrijk. De meeste talen hebben verschillende werkwoordstijden. Die geven aan of iets in het verleden, heden of de toekomst plaatsvindt en kunnen zelfs aangeven op welke manier het werkwoord het begrip van tijd veronderstelt.

Vendler (1967) stelt een schema voor dat als model van vergelijking kan dienen om werkwoorden en zinnen in te delen op basis van tijdsbegrip. Hij stelt dat er werkwoordsvormen zijn mét en zónder doorlopende tijd. Bij de zinnen mét doorlopende tijd zou je de vraag kunnen stellen *hoe lang iets duurt*. Het zijn processen in de tijd met opeenvolgende fases. Binnen deze groep werkwoorden maakt Vendler (1967) nog onderscheid tussen *accomplishment*-begrippen en *activity*-begrippen. Een voorbeeld van een accomplishment is “een cirkel tekenen”; daarmee moet je doorgaan totdat de cirkel af is. Als je stopt voordat de cirkel af is heb je geen cirkel getekend, maar zodra er een cirkel op papier staat was de afgelopen bezigheid wél het tekenen van die cirkel. Het halen van een bepaald doel heeft dus invloed op de beschrijving van de actie. Anders is dat met activiteiten, zoals bijvoorbeeld “rennen”; dat doe je voor een bepaalde tijd zonder dat het op een vastgesteld moment is afgelopen. Elk deel van het proces is van dezelfde aard als het geheel; de activiteit verloopt homogeen in de tijd.

Bij werkwoorden zónder doorlopende tijd gaat het om de vraag *op welk moment* iets gebeurt of *voor hoe lang iets het geval is*. Hier is de actie geen proces in de tijd dat in stapjes te verdelen is, maar het is wel een predicaat: de zin is op een bepaald moment waar of onwaar. Vendler (1967) maakt hierin onderscheid tussen *achievement*-begrippen en *state*-begrippen. Een voorbeeld van een achievement is “herkennen”; dat gebeurt op een bepaald moment. Het is een predicaat dat waar is

voor een enkel moment in de tijd. Een voorbeeld van een state is “weten”; je weet iets voor een korte of lange periode. Verder zit daar geen ontwikkeling in.

Het onderscheid tussen deze vier klassen lijkt voor zich te spreken. Toch is het gemakkelijk om achievements en accomplishments te verwarren. Zie bijvoorbeeld de zin “Hij vond de schatkist in 5 minuten“. Het lijkt een accomplishment met doorlopende tijd, omdat het 5 minuten duurde voor hij de kist vond. Het vinden zelf gebeurde echter maar op één moment. Activities en states zijn vrij makkelijk te onderscheiden. Zie bijvoorbeeld het verschil tussen “Ik ben vegetarisch aan het eten“ en “Ik eet vegetarisch“. Het eerste is duidelijk een activity, maar het tweede duidt eerder op een gewoonte. Gewoontes gedragen zich volgens Vendler (1967) meestal als states. Om dit te verduidelijken, maakt Vendler nog onderscheid tussen verschillende soorten states: specifieke en algemene states. Een voorbeeld van een specifieke state is het besturen van een taxi. Een taxichauffeur kan altijd zeggen dat hij een taxi bestuurt (dan is het een state), terwijl hij slechts soms werkelijk een taxi aan het besturen is (dan is het een activity). Een algemene state heeft die eigenschap niet; een voorbeeld is het heersen over een land. Een heerser kan altijd zeggen dat hij over het land heerst, maar is nooit werkelijk letterlijk aan het heersen op een specifiek moment. Een algemene state kan dus niet ook een activity zijn.

Tot zover hebben we dus een classificatie van werkwoorden in vier categorieën, met eventueel nog een extra onderverdeling van specifieke en algemene states. Zie tabel 1 voor een samenvattend overzicht van de vier klassen, een vrije interpretatie van het schema van Vendler (1967, p. 149).

Tabel 2: Vendler’s werkwoordclassificatie

Doorlopende tijd	Klasse	Voorbeeld	Omschrijving
Ja	Activity	Rennen	Proces; elk deel van dezelfde aard als het geheel.
	Accomplishment	Een cirkel tekenen	Proces met een bepaald eind.
Nee	Achievement	Herkennen	Predicaat dat waar kan zijn op een bepaald moment.
	State	Weten	Predicaat dat waar kan zijn gedurende een bepaalde tijd.

Is dit nu een volledige analyse van alle werkwoorden? Vendler (1967) zegt vrij bescheiden dat hij het slechts bedoeld heeft als model van vergelijking. Hij ziet zijn schema niet als dé waarheid en hij bedoelt het niet als een inkadering waar de realiteit in moet passen. Er is echter wel een enorm groot aantal werkwoorden dat volledig binnen één van de categorieën past. Vendler (1967, p. 149): “This division has an air of completeness about it. Perhaps it is more than a mere presumption to think that all verbs can be analyzed in terms of these four schemata.”

1.3.3 Doelgerichtheid en events

Als we het schema van Vendler (tabel 2) vergelijken met hoe Liem et al. de klassen van Vendler beschrijven (tabel 1), zien we een nuanceverschil. Vendler (1967) legt de nadruk op het verschil tussen processen en predicaten, tussen wel of geen doorlopende tijd. Liem et al. (2013) noemen juist de tweedeling tussen werkwoorden met en zonder hoogtepunt/eindpunt. In hun analyse van de tekstjes is juist het belang van dat verschil opgevallen.

Champollion & Krifka (2015) noemen dit verschil ook wel het onderscheid tussen telische (doelgerichte) en atelische (niet doelgerichte) werkwoorden. Ze analyseren taalkundige “onderdelen” op een wiskundige manier, met behulp van de verzamelingenleer. Werkwoorden zien zij als verzamelingen van *events* die verschillende

eigenschappen kunnen hebben, bijvoorbeeld het telisch of atelisch zijn. Maar wat zijn events dan precies?

In hetzelfde jaar dat Vendler zijn werkwoordclassificatie publiceerde, ontwikkelde Donald Davidson zijn semantische theorie op basis van de notie *event* (gebeurtenis) (Davidson, zoals geciteerd door Maienborn, 2010). Hij ziet events als concrete spatiotemporele objecten, oftewel als dingen met een locatie in ruimte en tijd. Hij stelt dat elk werkwoordpredicaat ook een verborgen event-argument heeft, dat voor de gebeurtenis zelf staat. Dan heeft in de zin “Jones buttered the toast” niet alleen Jones een relatie met de toast, maar omschrijft het een gebeurtenis - een plaats en tijd waarop dit plaatsvindt. Waar refereert het woordje “it” anders naar in de zin “Jones did it slowly”? Volgens Davidson wordt de wereld blijkbaar bevolkt door dingen als events; net als objecten kunnen die geteld worden, er kan naar gerefereerd worden, ze kunnen in ruimte en tijd gelocaliseerd worden en er kunnen meerdere eigenschappen aan toegeschreven worden.

Al vrij snel werd duidelijk dat event-argumenten veel wijder verspreid zijn en niet alleen maar gelden voor (actie-)werkwoorden (Maienborn, 2010). Zo kunnen bijvoorbeeld ook bijvoeglijk naamwoorden, zelfstandig naamwoorden en voorzetsels een event-argument hebben. Tegenwoordig worden event-argumenten dus gezien als een kenmerk van predicaten in het algemeen.

Het feit dat ze zo breed inzetbaar zijn, maakt het volgens Maienborn (2010) een logische stap om Davidson’s voorstel te combineren met Vendler’s (1967) classificatie van types van situaties; states, activities, accomplishments en achievements. De twee inzichten van Davidson en Vendler zijn een goede combinatie die de theorieën ook voor hedendaagse taalkundigen aantrekkelijk maken.

Zijn het slechts theorieën, of zit er meer achter? Hebben deze benaderingen ook een psychologische realiteit? Maienborn (2010) beschrijft een aantal psycholinguïstische studies die de taalkundige aannames over eventstructuur testen. Er is bijvoorbeeld onderzoek gedaan naar het verwerken (in de hersenen) van de onderliggende eventstructuur (Gennari & Poeppel, zoals geciteerd door Maienborn, 2010). Hieruit werd duidelijk dat het verwerken van zinnen met een event langer duurt dan het verwerken van statische zinnen. De onderzoekers stellen dat dit komt doordat werkwoorden waarbij iets gebeurt een complexere achterliggende structuur hebben. Maienborn (2010) ziet dit als empirisch bewijs voor het onderscheid tussen events en states, en als experimentele ondersteuning voor de psychologische realiteit van de compositionele structuur van natuurlijke taal. Hoewel het een controversieel onderwerp is en het onderzoeksveld van psycholinguïstisch onderzoek naar eventstructuur nog in de kinderschoenen staat, stelt Maienborn (2010) dat davidsoniaanse events een serieus onderwerp van psychologisch onderzoek worden. Er zijn steeds meer aanwijzingen dat verborgen event-argumenten behalve erg handig ook psychologisch waar zijn.

1.3.4 Toepassing door Liem et al. (2013)

Wat Liem et al. (2013) opviel aan de tekstjes was de schijnbare overeenstemming wat betreft doelgerichtheid. Ze hebben vier categorieën opgesteld, geïnspireerd op de klassen van Vendler (1967), maar dan met de nadruk op doelgerichtheid en de zekerheid van het behalen van een doel. Hun vier klassen zijn niet alleen toepasbaar op werkwoorden (zoals de klassen van Vendler), maar algemener op events/gebeurtenissen. Die gebeurtenissen kunnen ook een werkwoord of een enkele zin zijn, maar ze kunnen ook breder getrokken worden door een gehele tekst te bekijken. Dat is wat Liem et al. (2013) gedaan hebben: alle teksten bij alle fragmenten in hun geheel indelen in één van hun vier klassen. Zie tabel 3 voor de klassen die ze onderscheidde. De cursieve zinnen zijn voorbeelden, overgenomen uit de cinematografische beschrijvingen.

Tabel 3: De classificatie van Liem et al. (2013, p. 26)

Class 1	Activity or state with no goal. Vb: <i>“A group of people working in a factory. There is steam everywhere. They work hard, but they are not negative about that.”</i>
Class 2	Activity with underspecified goal and no certainty of achieving the goal. Vb: <i>“A scene of awaiting. Something is going to happen.”</i>
Class 3	Activity with well-specified goal and no certainty of achieving the goal. Vb: <i>“Warriors are moving fast in the darkness and trying to sneak into their enemies’ campsite.”</i>
Class 4	Activity with well-specified goal and certainty of achieving the goal. Vb: <i>“Someone is walking in a dense forest and finally arrives at village where the inhabitants are dancing slowly.”</i>

Van drie fragmenten waar veel beschrijvingen van waren, hebben Liem et al. (2013) alle beschrijvingen ingedeeld in één van de vier klassen. Bij twee van de drie fragmenten was er een duidelijke meerderheid bij een bepaalde klasse. Van het fragment *Exciting Trailer* werden bijvoorbeeld 57.94 % van de beschrijvingen ingedeeld bij klasse 3.

Verder hebben ze alle andere 54 fragmenten (met slechts enkele bijbehorende beschrijvingen) ook gecodeerd. Voor 44 van de 54 fragmenten viel er een meerderheid van de beschrijvingen in dezelfde klasse. Voor die gevallen hebben ze een zogenaamde “confusion matrix” opgesteld, waar ze een aantal dingen uit concludeerden. Ten eerste viel op dat wanneer er een meerderheid van de deelnemers een doelgerichtheid beschreef (klasse 2, 3 of 4), die meerderheid ook zeer groot was. Hetzelfde geldt voor de onduidelijkheid van de haalbaarheid van een doel (klasse 2 en 3). Ook daar waren de deelnemers het grofweg over eens. Klasse 1 (state of activity zonder doel) werd echter vaak gemixt met andere klassen.

We hebben nu de classificatie van Vendler (1967) gezien, de beschrijving die Liem et al. (2013) van Vendler gaven, en de alternatieve versie van Liem et al. In mijn onderzoek ga ik onderzoeken welke versie het best toepasbaar is op kleinere schaal; dus niet op hele teksten maar op enkele zinnen. In hoofdstuk 2 beschrijf ik hoe ik dat precies aanpak. Door meer in te zoomen op de teksten, hoop ik een genuanceerder verband te vinden tussen tekst en muziek.

1.4 Verbinding aan de muziek

1.4.1 Eigen redenen van deelnemers

Voor een eerste verbinding van tekst en muziek, vroegen Liem et al. (2013) hun deelnemers naar de motivatie voor hun beschrijving. Ze vroegen zich af of de deelnemers in staat waren aan te geven waarom de muziek hen voor een bepaalde scène liet kiezen. Daarmee waren ze op zoek naar narratieve elementen die schijnbaar een sterke (high-level) muzikale associatie hadden. Die high-level kenmerken zouden verbonden kunnen worden aan low-level aspecten die uit het signaal gehaald kunnen worden, en eventueel in de toekomst gebruikt worden als zoektermen in zoeksystemen.

Ze hebben alle beschrijvingen in twaalf categorieën ondergebracht, waarvan de eerste zes afkomstig zijn van de topologie van Tagg en Clarida (zoals geciteerd door Liem, Larson & Hanjalic, 2013). De overige zes categorieën hebben ze zelf opgesteld voor de motivaties die niet in de eerste zes categorieën pasten. De categorieën zijn: sonic anaphone, kinetic anaphone, tactile anaphone, genre synecdoche, episodic marker, style indicator, temporal development, character trait/psychology of actor, general quality/atmosphere, seen before, personal relation, intuition. Hierin

zagen ze drie opvallende metacategorieën:

1. Temporal aspects
(Ontwikkeling van tijd, beweging en veranderingen van episode.)
2. Psychology, quality and atmosphere
(Aanduiding van de psychologie van een karakter en opwekking van empathie.)
3. Previously seen examples
(Kennis van bestaande films en stijlen die de beschrijving gestuurd hebben.)

Liem et al. (2013) concluderen dat de deelnemers niet zo goed zijn in het uitleggen van hun connotatieve associaties in termen van muzikale kenmerken. Ze zijn wel gevoelig voor associaties met voorgaande metacategorieën. Ze laten inzicht zien in de relatie tussen temporele ontwikkeling en muzikale beweging, tussen affectieve en psychologische effecten en tonen kennis van bestaande voorbeelden.

Als het gaat om tijdsontwikkeling en beweging, lijken de kenmerken die associaties veroorzaken bij deelnemers zich op een hoger level te bevinden dan slechts de beat of het tempo, maar wel op een lager level dan formele muzikale vormanalyse. Wat betreft affect, blijkt de relatie tussen toonsoort en emotie ook genuanceerder te zijn dan de gangbare opvatting dat majeure toonsoorten positieve gevoelens opwekken en mineur toonsoorten negatieve. Zo waren er een aantal rustige stukjes in majeur die over het algemeen wel als positief beoordeeld werden, maar waarbij de gevoelens die beschreven werden in de scène wel gemengd waren. Ook waren er twee pianostukjes in mineur die over het algemeen juist als opgewekt en opzwevend ervaren werden.

Verder concluderen ze dat er bij de associaties van muziek naar tekst vaak stappen genomen worden die niets met de muziek op zich te maken hebben; bijvoorbeeld buitenmuzikale herinneringen aan persoonlijke situaties of associaties aan de hand van individuele luistergeschiedenis. Bovendien zijn associaties in hoge mate gebaseerd op culturele conventies.

Liem et al. (2013) noemen hun bevindingen momenteel nog anekdotisch, maar stellen wel dat ze in overeenstemming lijken met de theorieën over cognitie, verwachting en bekendheid die bijvoorbeeld door Huron (2007) zijn onderzocht.

1.4.2 Muzikale verwachting

David Huron (2007) beschrijft in zijn boek een psychologische theorie over muzikale verwachting. Muziek speelt met de verwachtingen van de luisteraar en wekt daarmee allerlei reacties en gevoelens op. Huron (2007) noemt drie elementen die meespelen in dit spel met verwachting: voorspelbaarheid, verrassing en spanning.

Hij begint met hoe die elementen biologisch gezien werken bij mensen. Het zijn biologische systemen met een bepaalde functie, zoals het voorbereiden op toekomstige situaties en het alert maken bij dreigende situaties. *Voorspelbaarheid* representeert een biologisch succesvol werkend systeem. Als we dingen goed kunnen voorspellen, hoeven we immers niet bang te zijn voor onverwacht gevaar. In de muziek betekent dat dat voorspelbaarheid kan leiden tot het *prediction effect*: de luisteraar ervaart plezier doordat zijn verwachtingen uitkomen. Als het te voorspelbaar wordt, is het natuurlijk saai en volgt een negatieve reactie.

Verrassing representeert de mogelijkheid van gevaar. Biologisch gezien komt er na verrassing een *fight, flight of freeze*-reactie, maar in muziek wordt de situatie onmiddellijk als niet-dreigend beoordeeld. De reacties die een muziekluisteraar ervaart bij verrassende muzikale gebeurtenissen zijn volgens Huron (2007) *huivering, gelach of ontzag*. Huron associeert deze reacties met het plezier dat komt kijken bij het luisteren naar muziek, door middel van het fenomeen *contrastive valence*. Als een positieve reactie een negatieve reactie opvolgt, wordt plezier verhoogd. Verrassing

is biologisch gezien iets slechts, maar doordat in muziek vrijwel altijd een positieve reactie volgt, is verrassing in muziek wel plezierig. Te veel verrassingen op een te onvoorspelbare manier kunnen natuurlijk ook negatieve gevoelens opwekken, door te onsamenhangend en irritant te zijn. Zonder contrastive valence is een verrassing gewoon raar of vervelend.

Met behulp van voorspelbaarheid en verrassing kan *spanning* opgebouwd worden. Als een toekomstige muzikale gebeurtenis erg waarschijnlijk/voorspelbaar is, leidt dat tot een gevoel van anticipatie: een sterk gevoel van de onvermijdelijkheid van die gebeurtenis. Die anticipatie leidt tot het ervaren van spanning - een spanning die proportioneel is met de voorspelbaarheid van de komende gebeurtenis.

1.4.3 Voorspelbaarheid

De basis van een samenhangend muziekstuk is de voorspelbaarheid. Een componist (bijvoorbeeld van filmmuziek) kan zich natuurlijk ook juist richten op angst of ongemak door de voorspelbaarheid expres op een vervelende manier te doorbreken. Bovendien kan te veel voorspelbaarheid juist leiden tot saaiheid. Over het algemeen is het correct voorspellen van bepaalde muzikale gebeurtenissen een plezierige ervaring. Er zijn een aantal manieren om die voorspelbaarheid te bereiken, maar hoe dat precies tot een luisterervaring leidt heeft ook te maken met de luisteraar zelf. Specifiek is het een combinatie van de dynamische werking van het geheugen (de psychologische structuur van het menselijke geheugen) en de specifieke luistergeschiedenis en muzikale blootstelling van de persoon zelf.

Huron (2007) noemt de volgende vier wegen naar voorspelbaarheid: schematische voorspelbaarheid, dynamische voorspelbaarheid, vertrouwdheid en bewuste voorspelbaarheid.

Schematische voorspelbaarheid ontstaat door het vasthouden aan bestaande schema's of normen die cultureel gedefiniëerd zijn, zoals het gebruik van bekende instrumenten, toonladders, maatsoorten, harmonieën en stijlen. Huron (2007) noemt de volgende voorbeelden:

- Melodische normen: central pitch tendency (gecentreerde/evenwichtige verdeling van tonen), pitch proximity (meer kleine dan grote intervallen), step declination (voornamelijk dalende lijnen).
- Ritmische normen: metric hierarchy, metric proximity (ritmes/maatsoorten wisselen niet te snel).
- Normen voor frases: frases met een boogvorm, Frases verdeeld over vier of acht maten.
- Harmonische normen: veelvoorkomende akkoordprogressies, stabiele harmonische ritmes, clichématige cadensen.

Dynamische voorspelbaarheid wordt bereikt door het vasthouden aan de verwachtingen/normen die in het stuk zelf gecreëerd zijn, zoals bijvoorbeeld een motief dat logisch klinkt in het geheel van het werk.

- Interne herhaling (op grote en kleine schaal)
- Ostinaten
- Sequensen
- Thema's en motieven, kenmerkend voor het werk

Vertrouwdheid wil zeggen dat het overeenstemt met de luisterervaring.

- Bijvoorbeeld door herhaalde blootstelling.

Bewuste voorspelbaarheid ontstaat door het gebruik van geijkte methoden, die een theoretisch onderlegde luisteraar bewust kan herkennen.

- Bijvoorbeeld wanneer een majeur stuk tijdelijk naar mineur moduleert en vervolgens weer teruggaat naar majeur.

1.4.4 Verrassing

Verrassing kan op twee manieren werken: vervaardigd tot genot (met behulp van contrastive valence), of juist opzettelijk ergerlijk. Verrassingen werken het beste als ze bewust opgezet zijn. Muzikanten wekken rillingen, gelach of ontzag op door de verwachtingen van de luisteraars te verstoren. De sociale/fysieke omgeving kan ook invloed uitoefenen op het effect van een verrassing. Andersom kan muziek juist ook zorgen voor of invloed hebben op een bepaalde sociale context.

Als het schenden van een bepaalde verwachting vaak gebeurt en dus gangbaar wordt, dan kunnen luisteraars leren om “het onverwachte te verwachten”. Een voorbeeld hiervan is de syncope: een accentverschuiving waar je, als je eraan went, tegelijkertijd verrast maar toch niet zo verrast door bent.

Huron (2007) noemt, net als bij het creëren van voorspelbaarheid, vier categorieën waarmee je verrassing kunt opwekken: schematische verrassingen, dynamische verrassingen, verrassing door bekendheid en bewuste verrassingen.

Schematische verrassingen ontstaan door het overtreden van een bestaand schema/een bestaande norm. Voorbeelden hiervan zijn:

- Veranderingen in dynamiek, instrumentatie, toonladders, maatsoorten, harmonieën, intervallen, stijlen, ...
- Plotselinge verandering van luidheid, plotselinge wisseling van instrumentatie, het overtreden van de huidige toonsoort, de maatsoort schenden, zeldzame akkoordopvolgingen gebruiken, vroegtijdig een frase beëindigen, een onwaarschijnlijke melodische sprong toevoegen, een oplossing uitstellen, naar een andere stijl springen, ...
- Bedrieglijke cadens: eindigen met de trappen V-vi in plaats van V-I. Dat is bedrieglijk omdat het minder voorkomt.
- Chromatische mediantprogressie: akkoorden die zich niet conformeren aan de toonsoort. Zijn zeer zeldzaam, dus erg verrassend. Worden vaak gebruikt in filmmuziek, waar ze geassocieerd worden met buitengewoon emotionele momenten.

Dynamische verrassingen komen door het overtreden van verwachtingen die opgewekt zijn in het stuk zelf. Dat kan een what- of een when-overtreding zijn. Het is moeilijk een strict, duidelijk voorbeeld van een dynamische verrassing te vinden, omdat ze vaak ook een element van schematische overtreding meebrengen. Enkele voorbeelden:

- Hemiool.
- Syncopen.
- “Garden-path-verrassing”: als eerdere delen van een frase, die al gehoord en begrepen waren, ineens opnieuw geanalyseerd moeten worden in het licht van een later deel. Een taalkundig voorbeeld is “the old man the boats”. Je zou in eerste instantie “the old man” als onderwerp begrijpen, tot het woordje “the” erachteraan komt en “man” een werkwoord blijkt te zijn. Hetzelfde komt in muziek voor, als bijvoorbeeld een opmaat verkeerd geïnterpreteerd wordt.

Verrassingen door *bekendheid* ontstaan doordat een stuk bekend geacht werd maar toch anders wordt uitgevoerd. Deze overtreding komt meestal voor bij activiteiten van de uitvoerders, niet van componisten. Voorbeelden zijn:

- Fout tijdens de uitvoering.
- Verkeerd citeren.

- Opzettelijke parodie.
- Nuance van de uitvoerder.

Bewuste verrassingen worden bereikt door het overtreden van een bewuste verwachting bij een goed geïnformeerde luisteraar.

- Voorbeelden hiervan staan ook bij de schematische verrassingen, die in dit geval heel bewust ervaren worden.

1.4.5 Spanning

Een derde element dat meespeelt bij de verwachtingen in muziek is spanning. Voorspelbaarheid zorgt voor de positieve *prediction response* als een muzikale gebeurtenis juist is voorspeld, maar hoe voorspelbaarder een toekomstige gebeurtenis is, hoe groter ook de *tension response* als het langer en langer duurt voor het gevoel van anticipatie daadwerkelijk bevredigd wordt. Dit mechanisme, anticipatie gevolgd door een positieve voorspellingsreactie, heet ook wel een *tension-and-release experience*. Verrassingen spelen hierin een grote rol door de *release* nog even uit te stellen. Een van de wezenlijkste middelen die hier gebruik van maken is de muzikale climax. Huron (2007) noemt een aantal manieren waarop spanning opgebouwd kan worden: spanning en oplossing, toenemende anticipatie en voorspellende spanning.

Spanning en oplossing kan worden opgewekt door middel van:

- Beweging van chromatische tonen naar aangrenzende diatonische tonen.
- Gangbare versieringen zoals voorhoudingen en anticipaties.
- Stereotiepe cadensen.

Toenemende anticipatie kan ontstaan door:

- Delay (vertraging/uitstel) bij het einde van een frase, bij de eerste tel van de maat (sterk maatdeel), op zwakke maatdelen, bij leidtonen of bij eindcadensen.
- Bij climaxen is er een grote verzameling mogelijkheden, zoals rubato, ritardando, accelerando en fade-out.

Voorspellende spanning is spanning die niet aan onmiddellijk aangrenzende gebeurtenissen gerelateerd is, maar aan een verder liggende gebeurtenis:

- Dit heeft alles te maken met het voorstellingsvermogen en de verbeelding.

1.4.6 Samenvattend schema muzikale verwachting

Tabel 4 geeft een samenvattend schema weer van wat Huron (2007) als voorbeelden en mogelijkheden noemt om te spelen met verwachting. In hoofdstuk 3 onderzoek ik of dit een volledige weergave is van de mogelijkheden voor het creëren van voorspelbaarheid, verrassing en spanning, en of dit verbonden kan worden aan de verhaalstructuur van zowel muziek als tekst.

De deelnemers van het onderzoek van Liem et al. (2013) konden hun associaties dan niet zo goed uitleggen in termen van muzikale kenmerken, wellicht hebben ze wel een aantal van deze elementen onbewust gehoord.

Tabel 4: Muzikale kenmerken verzameld uit Huron (2007)

Voorspelbaarheid	<p>Central pitch tendency (gecentreerde/evenwichtige verdeling van tonen) Pitch proximity (meer kleine dan grote intervallen) Step declination (voornamelijk dalende lijnen) Metrische hiërarchie Metric proximity (ritmes/maatsoorten wisselen niet te snel) Frases met een boogvorm Frases verdeeld over vier of acht maten Veelvoorkomende akkoordprogressies Stabiele harmonische ritmes Clichématige cadensen Interne herhaling (op grote en kleine schaal) Ostinaten Sequensen Thema's en motieven, kenmerkend voor het werk Herhaalde blootstelling</p>
Verrassing	<p>Plotselinge verandering van luidheid Plotselinge wisseling van instrumentatie Het overtreden van de huidige toonsoort De maatsoort schenden Zeldzame akkoordopeenvolgingen gebruiken Vroegtijdig een frase beëindigen Een onwaarschijnlijke melodische sprong toevoegen Een oplossing uitstellen Naar een andere stijl springen Bedrieglijke cadens Chromatische mediantprogressie Hemiool Syncopen “Garden-path”-verrassing Fout tijdens de uitvoering Verkeerd citeren Opzettelijke parodie Nuance van de uitvoerder</p>
Spanning	<p>Beweging van chromatische tonen naar aangrenzende diatonische tonen Gangbare versieringen zoals voorhoudingen en anticipaties Stereotiepe cadensen Vertraging/uitstel bij het einde van een frase Vertraging/uitstel bij de eerste tel van de maat (sterk maatdeel) Vertraging/uitstel op zwakke maatdelen Vertraging/uitstel bij leidtonen Vertraging/uitstel bij eindcadensen Rubato Ritardando en accelerando Fade-out Anticipatie op latere gebeurtenissen Voorstellingsvermogen</p>

2 Methode

2.1 Algemeen

Mijn onderzoek bouwt voort op het onderzoek van Liem, Larson & Hanjalic (2013). Ik onderzoek de relatie tussen een muziekfragment en cinematografische beschrijvingen van filmscènes die bij die muziek zouden kunnen passen. De uitkomst zou mogelijk kunnen bijdragen aan onderzoek naar de mogelijkheden voor het karakteriseren van muziek in termen van verhaalstructuur.

Ik gebruik de beschrijvingen van één muziekfragment (*Exciting Trailer*) en vergelijk hun onderliggende verhaalstructuur. Ik ga op zoek naar overeenkomst, om vervolgens twee vragen te beantwoorden:

1. Is er een “gemiddelde” onderliggende verhaalstructuur in de cinematografische beschrijvingen van het stuk?
2. Kan de tekstuele verhaalstructuur verbonden worden aan specifieke momenten in de muziek?

Voor het beantwoorden van vraag 1 kijk ik onder andere de volgorde van klassen, naar segmentatie en naar dominante klassen (voor verdere toelichting, zie paragraaf 3.1). Een gemiddelde structuur zou positief zijn voor de verdere analyse, omdat het aannemelijk wordt dat die gemiddelde verhaalstructuur één op één samenhangt met de structuur in de muziek. Dat zou mogelijkheden bieden voor de verbinding van algemene narratieve elementen aan muzikale kenmerken. Mocht er géén gemiddelde structuur blijken te zijn, is het lastiger om specifieke tekstuele elementen aan muziek te koppelen, maar zal een ander soort vergelijking gemaakt moeten worden.

Het zoeken naar een antwoord op vraag 1 voer ik uit in twee fases: een testfase met één stukje tekst en een officiële fase met meerdere stukjes tekst. In de testfase experimenteer ik met het analyseren van één tekstje, *zin voor zin*. Ik probeer eerst de classificatie van Vendler (1967) uit, en bekijk welke problemen daarbij komen kijken. Vervolgens test ik de classificatie die Liem, Larson & Hanjalic (2013) gebruikten voor hele teksten, maar pas hem nu toe per zin. Vervolgens kies ik de beste optie en optimaliseer hem eventueel. In de officiële fase pas ik vervolgens die geoptimaliseerde classificatie toe op meerdere beschrijvingen. Zo kan ik een aantal gevonden structuren met elkaar vergelijken en ontdek ik of er overeenkomsten zijn, die mogelijk duiden op een vaste, globale onderliggende structuur.

Voor vraag 2 zoek ik de samenhang tussen tekst en muziek, door eerst het fragment op gehoor in te delen in blokjes, dat vervolgens te koppelen aan high-level kenmerken van Huron (2007) en te kijken of dit ook zichtbaar is in low-level visuele analyses van de audio. Vervolgens probeer ik voorzichtige conclusies te trekken over hoe Huron aan de gebruikte classificatie gekoppeld kan worden en welke eigenschappen van Huron (2007) veelbelovend zijn voor vervolgonderzoek.

2.2 Onderzoeksgroep

De onderzoeksgroep is gelijk aan die van Liem, Larson & Hanjalic (2013). Zij gebruikten het *Amazon Mechanical Turk* crowdsourcing platform, dat gemotiveerde deelnemers betaalt voor elke goedgekeurde reactie. Liem et al. motiveren uitgebreid waarom dit platform een betrouwbare methode is voor het verkrijgen van respons; ze noemen onder andere het grote aantal deelnemers, dat deelnemers er geld voor krijgen en dat slechte reacties geweigerd konden worden.

Per deelnemer is de volgende informatie bekend: geslacht, leeftijd, land van herkomst, opleiding, wel of geen muziekluisteraar, kijkt wel of geen webvideo's, kijkt wel of geen films. Voor het doeleinde van dit onderzoek is dit echter niet zo relevant.

Voor het fragment dat ik heb uitgekozen, hebben 108 deelnemers een cinematografische beschrijving gemaakt.

2.3 Meetinstrumenten

2.3.1 Classificaties en fragmenten

In stap 1 analyseer ik de tekstjes met cinematografische beschrijving. Voor de testfase gebruik ik in eerste instantie de classificatie van Vendler (1967) zoals ik die in het voorgaande hoofdstuk besproken heb, en de classificatie van Liem, Larson & Hanjalic (2013). Na de testfase kies ik één van de classificaties en pas ik die zonedig aan. Voor de volledigheid geef ik de initiële classificaties hier nog eens weer:

Vendler:

- A State: predicaat dat waar kan zijn gedurende een bepaalde tijd (*weten*).
- B Achievement: predicaat dat waar kan zijn op een bepaald moment (*herkennen*).
- C Activity: proces, waarbij elk deel van dezelfde aard is als het geheel (*rennen*).
- D Accomplishment: proces met een bepaald eind (*een cirkel tekenen*).

Liem, Larson & Hanjalic:

- 1 Activity or state with no goal
- 2 Activity with underspecified goal and no certainty of achieving the goal.
- 3 Activity with well-specified goal and no certainty of achieving the goal.
- 4 Activity with well-specified goal and certainty of achieving the goal.

Ik heb gekozen voor het fragment *Exciting Trailer* van Kevin MacLeod. Daarvan zijn 108 zeer diverse beschrijvingen. De instructies die Liem et al. (2013) gaven: *describe an idealized cinematic situation, for which a given music fragment would fit as a soundtrack*. Verder gaven ze de respondenten alle vrijheid.

De beschrijving van de componist is: *“A militaristic snare drum march begins this piece, reminiscent of Eastern Europe during World War II. The orchestral sounds blend with the sharpness of the brass throughout, but especially after 0:36, when the music is lyrical and melodic, though intensely emotional and suspenseful. At 0:52, the rhythmic march begins again, with the strings playing the melody and the brass reaching discordant, swelling heights until the final crescendo at 1:10.”*

Niet alle cinematografische beschrijvingen zijn even bruikbaar. Ik verwijder van te voren de te korte tekstjes of tekstjes die “niet in de juiste vorm” geschreven zijn. Bijvoorbeeld:

- “This is spy music!” - deze tekst is simpelweg te kort om te analyseren.
- “I would use this music fragment to depict rise of a powerful kingdom” - deze tekst is, naast te kort, moeilijk toepasbaar op de classificaties, omdat de ik-vorm gebruikt wordt.

Wat ik wel zie als “de juiste vorm” zijn teksten die stap voor stap beschrijven wat er gebeurt, niet teveel verwijzen naar cameratechnische aanwijzingen en de ik-vorm vermijden, zoals bijvoorbeeld:

- “The mistress of a prominent king is running through a dark, damp and dreary forest. She is in search of her true love [...]”
- “A warrior is preparing himself for battle. He is putting on his armour making his weapons ready [...]”

Ik heb simpelweg de beschrijvingen met de meeste narratieve informatie en de duidelijkste vorm gekozen. Zie bijlage A voor de beschrijving die ik heb gebruikt voor de testfase, de tien beschrijvingen van de officiële fase en de rest van de beschrijvingen, die ik niet heb geanalyseerd.

2.3.2 Audio-analyse

In stap 2 analyseer ik het audiofragment eerst op gehoor. Volgens onderzoek ik aan de hand van Huron (2007) en een visuele audio-analyse of mijn eigen analyse adequaat is. Ten eerste koppel ik mijn bevindingen aan Huron's high-level muzikale eigenschappen op het gebied van voorspelbaarheid, verrassing en spanning. Ten tweede toon ik wat low-level visuele analyses, verkregen met behulp van het programma *Sonic Visualizer*. Tot slot probeer ik een voorzichtige koppeling te maken tussen de audio-analyse en de cinematografische beschrijvingen.

2.4 Procedure

2.4.1 Stap 1: onderliggende verhaalstructuur

In de testfase ben ik in eerste instantie begonnen met het analyseren van één tekst: *"The mistress of a prominent king is running through a dark, damp and dreary forest. She is in search of her true love, a love that was forbidden. He was the heir to the rival kingdom and had been taken away to be killed for sleeping with the king's mistress. She had gotten word that he managed to escape so she fled the palace in search for him. She's running as fast as her feet will carry her, the tight bodice of her Victorian dress is making it increasingly hard for her to breathe, she is nearly gasping for air. She knows the king's men aren't far behind her, she can hear them calling out to her and the horses hooves clapping on the ground. As she pushes deeper into the forest it becomes almost impossible to walk in let alone run. Just then she trips on root as she falls to the ground thoughts on failure run through her she lets out an almost silent "I'm sorry my love I tried and failed you", just as her fall is broken, she's sure it is the king's men...as she raises her toward her capture she is met with the dark dangerous eyes of her love, he says to her "you have not failed" as he brings her to her feet and kisses her deeply. "Hurry we must continue on", he tells her. They begin to run forcing their way through trees, vines and thick vegetation however it isn't long before the king's men catch up and surround them."*

Deze beschrijving heb ik geanalyseerd aan de hand van de beide classificaties: zowel de klassen van Vendler (1967) als de vier categorieën die Liem et al. (2013) gebruikten. Zie bijlage A voor alle teksten en bijlage B voor alle analyses.

Ik ben begonnen met Vendler. Ik koos ervoor om alleen de werkwoorden te analyseren die betrekking hebben op de vertelde tijd; of in ieder geval woorden die zichtbaar zouden kunnen zijn in de beelden. De rest zie ik als achtergrondinformatie die uiteindelijk niet aan specifieke momenten uit de verhaalstructuur gekoppeld kan worden. Bepaalde sferische informatie zou natuurlijk wel aan sferische elementen van de muziek gebonden kunnen worden (bijvoorbeeld door "in a damp and dreary forest" aan donkere klanken te verbinden). Het gaat me hier echter specifiek om verhaalstructuur, dus negeer ik dat soort woorden en focus alleen op werkwoorden die in de beelden terug zouden komen. Ook werkwoorden die bepaalde zaken beschrijven die niet op het moment zelf gebeuren negeer ik. Een voorbeeld daarvan is "She's searching for her true love, a love that **was** forbidden". Het woordje "was" is wel een werkwoord, maar zegt alleen indirect iets over wat er op dat moment gebeurt. Het heeft dus niets met de verhaalstructuur van de beelden te maken. Tot slot negeer ik werkwoorden die zich intern bij een karakter afspelen, in hun hoofd. Die zijn ook moeilijk aan de beelden te koppelen zonder al te veel te speculeren. Een voorbeeld hiervan zou zijn "She's sure it's the king's men".

Wat de Vendlerclassificatie lastig maakt, zijn bijwoorden die de zin ontkrachten, zoals "bijna". Zie bijvoorbeeld de zin "She's nearly gasping for air". Als we gewoon alleen het werkwoord classificeren, krijgen we een vertekend beeld van wat er bedoeld wordt. Dit soort woorden geeft een andere dimensie aan het verhaal, maar zijn lastig in de werkwoordanalyse te vatten.

Al snel bleken de klassen van Vendler vrij nietszeggend voor het beoogde doel. Zo zijn erg veel werkwoorden (met name de werkwoorden die interessante wendingen aan het verhaal geven) geclassificeerd als klasse C (activity), maar dat zegt verder niets over of er een doel inzit of niet. Zo zijn bijvoorbeeld de volgende zinnen allebei van klasse C:

- The mistress is running.
- She is in search of her true love.

Dat zijn allebei activities – processen met doorlopende tijd. We kunnen ons afvragen: “hoe lang duurde het”. Het zijn geen states of achievements, want er zit wél ontwikkeling in (het is geen state zoals weten - dat is gewoon waar of onwaar gedurende bepaalde tijd). Ook zijn het geen accomplishments, want er zit geen einde aan. “Being in search of” hangt niet af van het vinden. Ook als ze niet vindt, was ze nog steeds aan het zoeken. Het zijn dus allebei activities, ook al zien we intuïtief een verschil tussen de twee. Rennen is gewoon rennen, maar zoeken heeft wel een doel, of je dat doel nu haalt of niet. Wellicht zijn de categorieën van Liem et al. (2013) instructiever.

De classificatie van Liem et al. (2013) bleek lastig toe te passen op werkwoorden. In plaats daarvan keek ik naar hele zinnen. Dit had wellicht ook gedaan kunnen worden met de Vendlerclassificatie; het bekijken van Vendler op hogere abstractie-niveaus dan werkwoorden had misschien iets interessanter opgeleverd. Ik heb dit echter niet meer geprobeerd, omdat de afwezigheid van doelgerichtheid bij Vendler voldoende reden was om de classificatie niet te kiezen. Verhaalstructuur en muziek draaien beiden om spanning en het wel of niet behalen van een eindpunt/verlossing van die spanning.

Bij de Liemclassificatie kijk ik dus in plaats van alleen naar het werkwoord naar de hele zin. Verder werk ik met dezelfde aanpak: ik analyseer weer alleen zinnen in vertelde tijd die in de beelden naar voren zouden kunnen komen. Een bijkomend voordeel van het kijken naar hele zinnen is dat voorkomen wordt dat het woordje “het” een gekke invloed uitoefent. Zie bijvoorbeeld de zin “The tight bodice of her Victorian dress is making it increasingly hard for her to breathe.” Als het onderwerp “the dress” is, dan probeert die jurk het ademen moeilijk te maken en slaagt daar ook in (klasse 4, activity with well-specified goal and certainty of achieving the goal). Maar wat met deze zin bedoeld wordt is dat “ze” moeilijk kan ademen door haar jurk, zonder verder doel in het verhaal (activity or state with no goal). Ik kies ervoor om in dit soort gevallen naar het geheel te kijken – in dit geval zou ik dus klasse 1 kiezen. Dit is een vrij rigoreuze keuze; het kijken naar de plaats van de zin in het geheel in plaats van naar de ontwikkeling binnen de zin.

Het kijken naar hele zinnen in plaats van enkel naar werkwoorden heeft als bijkomstigheid dat het woordje bijna in “She is nearly gasping for air” geen moeilijkheden meer oplevert. Het werkwoord op zichzelf zou een klasse 4 moeten zijn (activity with well-specified goal and certainty of achieving the goal), als je het zo ziet dat ze naar lucht hapt met als doel lucht te krijgen en daarin dus slaagt. Echter, het woordje bijna maakt het verschil. Het lijkt daardoor meer een “activity or state with no goal”, omdat ze niet écht naar lucht hapt en daardoor ook niet ergens in kan slagen. Door het woordje bijna staat de hele zin in een ander licht. In dit soort gevallen kies ik dan ook voor klasse 1 (activity or state with no goal).

Het voorgaande toont dus goed het verschil tussen Vendler en Liem et al. Vendler zou hier een D van maken (accomplishment, proces met een bepaald eind), omdat het werkwoord nu eenmaal in die categorie hoort, maar Liem’s klasse 1 (activity or state with no goal) past veel beter bij de zin, wat die uitdrukt in het geheel en wat er waarschijnlijk op dat moment in de muziek gebeurt.

Terwijl ik de klassen aan het toewijzen ben, kom ik erachter dat er nog meer nuance gemaakt kan worden. Er is een uitbreiding mogelijk die net iets meer zou

kunnen zeggen over de structuur in de muziek en de spanningsopbouw wat betreft doelmatigheid. Ik voeg daarom twee klassen toe aan de classificatie:

5 Not achieving the goal.

6 Achieving the goal.

Het behalen van een doel zou vaak geassocieerd worden als klasse 4 (activity with well-specified goal and certainty of achieving the goal), maar het geeft veel meer informatie als weergegeven wordt dat het doel inderdaad behaald is. Dit kan waarschijnlijk ook goed gekoppeld worden aan bijvoorbeeld een eindcadens in de muziek. Op dezelfde manier geeft het juist niet behalen van een doel ook veel informatie over de spanningsopbouw.

Neem bijvoorbeeld de zin “She falls to the ground”. Dit zou een klasse 4 kunnen zijn, want ze valt; het doel is het raken van de grond en dat doel bereikt ze ook. Maar het kan ook een 1 zijn, want het op de grond vallen heeft verder geen hoger doel. Ik zou het eigenlijk liever zien in het licht van het geheel. Ze is aan het vluchten om haar achtervolgers te verliezen, maar daarin faalt ze (tijdelijk) door op de grond te vallen. Ik maak er dus een klasse 5 van (not achieving the goal), omdat dat (met enige mate van speculatie) het best correspondeert met wat de luisteraar in de muziek hoorde.

Met deze testfase heb ik bereikt dat ik een opeenvolging heb van Vendlerklassen (CCDCCDACBDDCDBCDDCDD) en een opeenvolging van Liemklassen (236211133551222114633355) - zie wederom bijlage B. Deze staan echter nog niet volledig op volgorde van de beelden. Er is nog een laatste stap nodig in de analyses: chronologisering. De volgorde van zinnen is namelijk niet noodzakelijk een letterlijke chronologische weerspiegeling van de scene. Daar werd door Liem et al. (2013) ook niet naar gevraagd. Ik pas een lichte vorm van chronologisering toe, door zinnen die duidelijk iets zeggen over het verleden, of een verklaring zijn voor acties die nu gebeuren, naar voren te halen in de analyse. Bijvoorbeeld de zin “she fled the palace in search for him”: dat is de reden dat ze nu in het bos loopt. Ze loopt niet éérs in het bos om daarna uit het paleis te vluchten. Hiermee los ik echter niet alles op, dus voor ik conclusies kan trekken zal ik hier nog verder over na moeten denken. Na het op de juiste volgorde zetten van de klassen, krijgen we de opeenvolgingen zoals te zien in Tabel 5:

Tabel 5: User Description I

V:	D	C	C	C	C	D	A	C	B	D	D	C
C:	6	2	3	2	1	1	1	3	3	5	5	1

D	B	C	B	D	D	C	D	D	C	D	D
2	2	2	1	1	4	6	3	3	3	5	5

Zoals ik al zei is de Vendlerclassificatie (de eerste rij van de tabel) bij nader inzien toch niet zo handig, omdat het, vooral als het een C-klasse betreft, weinig zegt over de doelmatigheid van de zin. In de officiële fase zal ik de Vendlerklassen dan ook verwerpen en kies ik volledig voor de uitgebreide versie van Liem et al. (2007), met toegevoegde klasse 5 en 6.

De tweede regel van de tabel geeft de Liemclassificatie weer, waarin duidelijk te zien is dat de reeks begint met het behalen van een doel (uit een kasteel breken), vervolgens met een aantal 1-, 2-, en 3-klassen toewerkt naar een bepaald doel (namelijk het vinden van haar liefde), vervolgens daarin faalt met twee 5-klassen (ze valt en denkt verloren te zijn), dan weer een aantal 1-, 2- en 3-klassen doorloopt om uiteindelijk tóch te slagen (haar geliefde blijkt haar gevonden te hebben), dan weer

met wat 3-klassen verder te vluchten om uiteindelijk tóch te falen met 5-klassen (weer ingehaald worden door de vijand).

Met behulp van deze classificatie is de verhaalstructuur dus makkelijk in te delen in stukjes die naar een 5- (faal-) of een 6- (slaag-) klasse toeleiden. Dat zouden we kunnen gebruiken als een manier van segmentatie. Deze methode zal ik ook in de officiële fase toepassen. Als we deze segmentatie visualiseren wordt de structuur zoals in Tabel 3:

Tabel 6: User Description I - Liemclassificatie chronologisch

6	2 3 2 1 1 1 3 3 5 5	1 2 2 2 1 1 4 6	3 3 3 5 5
---	---------------------	-----------------	-----------

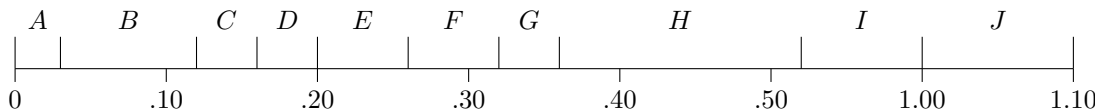
Zonder de muziek gehoord te hebben, zou iemand zich nu kunnen voorstellen dat dit in de muziek klinkt als een korte intro en vervolgens drie keer een opbouw met eerst een onbevredigend eind, vervolgens een bevredigend eind, maar daarna weer een negatieve climax (al is dit natuurlijk speculatie).

Met voorgaande ervaringen ben ik begonnen aan de officiële fase. Hierin heb ik nog eens tien beschrijvingen geanalyseerd. Ik heb dezelfde keuzes gemaakt voor de classificatie, ik heb dezelfde kleinschalige chronologisering toegepast en ik heb wederom geprobeerd een segmentatie te maken. Voor de segmentatie heb ik wederom het bereiken van een klasse 5 of 6 als basis gebruikt. Op deze manier ben ik tot elf analyses gekomen (de analyse uit de testfase ook meerekenend), die ik in paragraaf 3.1 zal tonen en nader zal bestuderen.

2.4.2 Stap 2: koppeling tekst en muziek

Bij stap 2 - het analyseren van de audio en een connectie zoeken met de beschrijvingen - ben ik begonnen met een high-level aanpak, om die vervolgens kort te vergelijken met low-level visuele audio-analyses. Ik begon intuïtief, om mijn eigen vondsten vervolgens met de verwachtingstheorie van Huron (2007) te vergelijken. Zo ontdekte ik een aantal veelbelovende elementen van Huron, die interessant zouden kunnen zijn om uit het audiosignaal te kunnen abstraheren. Mijn visuele analyses zijn slechts een eerste poging om te zien of mijn bevindingen inderdaad terug te zien zouden kunnen zijn.

Ik ben begonnen met zelf aandachtig luisteren naar het fragment. Op basis van mijn eigen gehoor en muzikale kennis, heb ik het fragment in zo veel mogelijk blokjes opgedeeld. Vervolgens heb ik gekeken naar de segmentatie in de tekstjes en voor elk voorkomend aantal segmenten een verdeling voorgesteld. Voor een eerste indruk, hieronder alvast mijn maximale analyse van het fragment.



In paragraaf 3.2 en verder zal ik dit uitgebreider beschrijven en analyseren. Ook toon ik daar een schematische samenvatting van de elementen die Huron (2007) noemt op het gebied van voorspelbaarheid, verrassing en spanning, om die gemakkelijker te kunnen vergelijken met mijn analyses. Ook zijn daar de low-level visualisaties en verdere analyses van de audio te vinden, en wat ik heb gevonden bij het vergelijken van tekst en muziek.

3 Analyse

3.1 Tekstanalyse

Zoals in paragraaf 2.4 besproken, heb ik de beschrijvingen zin voor zin geanalyseerd aan de hand van een aangepaste versie van de classificatie van Liem et al. (2013).

Voor de volledigheid hier nog eens het overzicht van de gebruikte klassen:

- 1 Activity or state with no goal
- 2 Activity with underspecified goal and no certainty of achieving the goal
- 3 Activity with well-specified goal and no certainty of achieving the goal
- 4 Activity with well-specified goal and certainty of achieving the goal
- 5 Not achieving the goal (whether as a surprise or not)
- 6 Achieving the goal (whether as a surprise or not)

Vervolgens heb ik de gevonden reeks klassen in segmenten verdeeld volgens de besproken methode. Zie bijlage B voor de volledige analyses, en hieronder het resultaat in tabellen:

Tabel 7: User Description I

6	2 3 2 1 1 1 3 3 5 5	1 2 2 2 1 1 4 6	3 3 3 5 5
---	---------------------	-----------------	-----------

Tabel 8: User Description II

1 1 3 2 1 1 4 4 6 6

Tabel 9: User Description III

6	3 3 2 6
---	---------

Tabel 10: User Description IV

3 3 2 3 2 4 6	3 3 3
---------------	-------

Tabel 11: User Description V

4 4 4 4 6	1 4 4
-----------	-------

Tabel 12: User Description VI

3 3 3 1 1 1 1 1 1 1 1 3 1 1 6

Tabel 13: User Description VII

3 3 3 1 3 5	3 3 1 3
-------------	---------

Tabel 14: User Description VIII

5	3 5	1 1 2 1 1 1 1 1 4 4 6	2 2
---	-----	-----------------------	-----

Tabel 15: User Description IX

4	1	1	1	1	4	3
---	---	---	---	---	---	---

Tabel 16: User Description X

2	1	2	4	1	4	1	1	1	1	1	6
---	---	---	---	---	---	---	---	---	---	---	---

Tabel 17: User Description XI

2	1	2	1	6	2
---	---	---	---	---	---

Vooraf stelde ik mezelf de volgende vraag: *Is er een “gemiddelde”, onderliggende verhaalstructuur in de cinematografische beschrijvingen van het stuk?*

Voor ik die vraag beantwoord, wil ik eerst wat dieper ingaan op het probleem van de volgorde/chronologie. Er werd door Liem et al. (2013) nu eenmaal niet gevraagd naar tekst met een letterlijke chronologische weerspiegeling van de cinematografische beelden. Niet elke zin heeft dus noodzakelijk een parallel in de muziek. Ik heb wel een simpele vorm van chronologisering toegepast, maar is dat genoeg? Taalkundige nuances zorgen ervoor dat soms een situatie op kleine schaal beschreven wordt met onduidelijkheden over wat zich in welke volgorde afspeelt. Zie bijvoorbeeld de zinnen “The king has a large army. He is marching towards a village.” De eerste zin heeft geen doelgerichtheid (klasse 1), en de tweede zin wel (klasse 4). Als de twee zinnen omgedraaid worden, blijft de betekenis hetzelfde: “The king is marching towards a village. He has a large army.” Nu is de volgorde echter van 1-4 veranderd naar 4-1.

Dit soort volgordeverschillen zijn moeilijk chronologisch te krijgen, simpelweg omdat chronologie geen rol speelt in dit soort gevallen. Het is slechts een beschrijving van de situatie en speelt zich tegelijkertijd af. De connectie van klassen 1, 2, 3 en 4 met de muziek is daarom moeilijk te maken. Gelukkig zijn de klassen 5 en 6 wel betrouwbaar als het gaat om volgorde. Het is over het algemeen duidelijk welke gebeurtenissen zich afspelen voor en na het wel/niet behalen van een doel. Dat maakt dat de segmentaties in blokjes die naar een 5 of een 6 toewerken inzichtelijker dan de volgorde van klassen op zichzelf.

Naast chronologisering, is het nodig nader in te gaan op de definitie van “gemiddelde” verhaalstructuur. Wat zou het precies kunnen inhouden en hoe zou je het kunnen vinden. Ik stel de volgende mogelijkheden voor als opties van waar je naar zou kunnen kijken bij het vergelijken van reeksen klassen:

- Klassenvolgorde
- Segmentatie
- Dominante klassen

In eerste instantie dacht ik aan exacte reeksen van klassen, maar de problemen met verschillen in lengte en onduidelijke chronologie maakt dit onmiddellijk een onaantrekkelijke optie.

Een betere optie is het kijken naar de segmentatie. Bij alle scenes is duidelijk te zien óf en hoe vaak er naar een doel toegewerkt wordt, of dat doel gehaald wordt of niet en of de scene op andere manieren duidelijk een overgang heeft naar een volgend segment. Als we de tabellen bekijken, zien we de volgende verschillen in aantal segmenten: 5, 3, 2, 2, 2, 1, 2, 4, 1, 1, 2. Meest voorkomend is dus een verdeling in twee segmenten. Het is echter gevaarlijk om aan te nemen dat de luisteraars ook echt niet meer dan twee onderdelen hebben gehoord in de muziek. Bovendien verschillen de tabellen met een twee-verdeling onderling ook nog behoorlijk: vier

tabellen gaan in het eerste deel naar een 6 (doel behaald), maar de vijfde naar een 5 (doel niet behaald). Bovendien is bij de ene de opbouw heel kort (meteen een 6 in het begin), en bij anderen weer langer. Ook het einde verschilt: in sommige tabellen eindigt het stuk op een 6, bij anderen werkt het tweede deel helemaal niet meer naar een doel toe.

Segmentatie blijkt een goede manier te zijn om te bepalen of er een onderliggende structuur bestaat, omdat het geen last heeft van verschil in lengte en duidelijk structuur weergeeft. Het antwoord op de vraag is in dit geval echter: nee, de tekstjes hebben geen gemiddelde onderliggende structuur wat betreft segmentatie. Daarvoor zijn de analyses te verschillend.

Een laatste manier van reeksvergelijking is het vinden van dominante klassen. Hieronder een tabel met de voorkomens van klassen per tabel:

Tabel 18: Voorkomens van klassen

	Klasse 1	Klasse 2	Klasse 3	Klasse 4	Klasse 5	Klasse 6
Beschrijving I	6	5	6	1	4	2
Beschrijving II	4	1	1	2	0	2
Beschrijving III	0	1	2	0	0	2
Beschrijving IV	0	2	6	1	0	1
Beschrijving V	1	0	0	6	0	1
Beschrijving VI	10	0	4	0	0	1
Beschrijving VII	2	0	7	0	1	0
Beschrijving VIII	7	3	1	2	2	1
Beschrijving IX	4	0	1	2	0	0
Beschrijving X	7	2	0	2	0	1
Beschrijving XI	2	3	0	0	0	1

Wat opvalt is dat er in enkele gevallen een vrij evenredige verdeling is van de klassen (bijvoorbeeld beschrijving I), en in andere gevallen juist niet (bijvoorbeeld beschrijving VI). De klassen die het vaakst buitenproportioneel voorkomen zijn 1, 3 en 4. In de volgende tabel geef ik weer wat de dominante klassen zijn en wat de originele classificatie was die Liem et al. (2014) gaven:

Tabel 19: Dominante klassen en Liemclassificatie

	Dominante klasse(n)	Originele klasse van Liem et al (2013)
Beschrijving I	1 en 3	3
Beschrijving II	1	4
Beschrijving III	3 en 6	4
Beschrijving IV	3	3
Beschrijving V	4	4
Beschrijving VI	1	4
Beschrijving VII	3	3
Beschrijving VIII	1	3
Beschrijving IX	1	3
Beschrijving X	1	4
Beschrijving XI	2	2

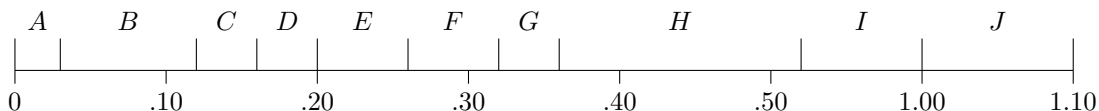
Blijkbaar is er geen directe samenhang met de dominante klassen en de originele classificatie van Liem et al. (2013). Dat is ook niet zo gek, want zij gebruikten maar vier klassen in tegenstelling tot zes. Maar ook de klassen die onveranderd zijn gebleven (1 t/m 4) hebben vaak een andere overheersende klasse in de zin-voor-zin-

analyse dan het label dat ze op de hele tekst kregen. Zie bijvoorbeeld beschrijvingen II, VI en VIII-X.

De aanpak van het kijken naar dominante klassen levert ons wederom de conclusie op dat er geen gemiddelde verhaalstructuur te vinden is in deze tekstuele beschrijvingen. Toch ga ik in de volgende paragrafen kijken of er een koppeling te maken valt tussen tekst en muziek.

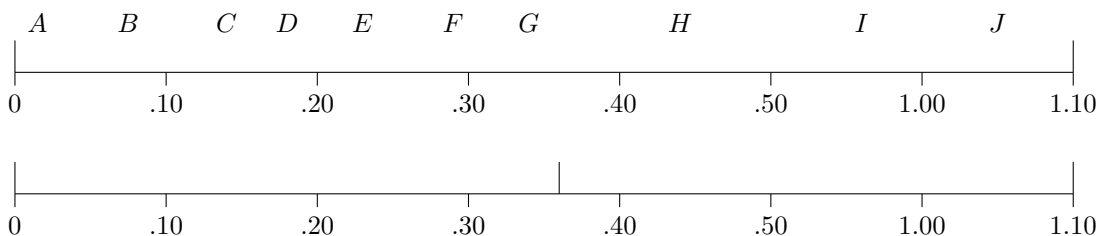
3.2 High-level audio-analyse

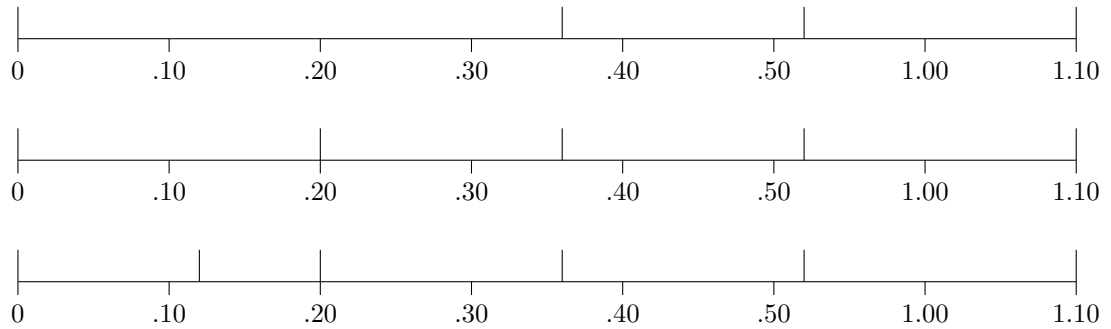
Voor stap 2 - het analyseren van de audio en een koppeling maken tussen de beschrijvingen en het muziekfragment - heb ik eerst zelf naar het fragment geluisterd. Op basis van mijn eigen gehoor en muzikale kennis, heb ik het fragment in zo veel mogelijk blokjes opgedeeld. Het resultaat is een tijdlijn van 1.10 minuut, verdeeld in tien blokjes, A t/m J. Hieronder deze maximale analyse, met eronder per blokje een beschrijving van de muziek:



- A* Alleen bas en percussie als een soort introductie.
- B* Strijkerslijn erbij in lange noten, blazers erbij met kort herhalend motiefje.
- C* Meer beweging en een dissonante toon in de strijkerslijn.
- D* Blazers vallen weg en een klokkenspel klinkt. Aan het eind een kleine crescendo.
- E* Strijkers een tweestemmig “dreigend” herhalend motief. Blazers zijn nog steeds weg, maar percussie loopt nog door.
- F* “Statige” blazersmelodie komt erbij.
- G* Blazers maken dissonante akkoorden. Crescendo naar een plotselinge overgang.
- H* Heel andere sfeer: drum valt helemaal weg, blazers en strijkers maken lange legato noten met hier en daar melodie. Begint lieflijk maar er komt dissonantie bij.
- I* Na een kort blazersmotiefje valt de muziek weer ongeveer terug op blokje E: drums vallen weer in en de strijkers spelen weer het herhalende motiefje.
- J* Blazers komen erbij en maken een grote opbouw naar het eind: twee keer een motief dat dissonant eindigt, met een grote crescendo eindigend in een opgelost akkoord, ritmisch en bombastisch gespeeld door de blazers.

Vervolgens heb ik de meest waarschijnlijke tijdlijnen opgesteld voor een verdeling in respectievelijk 1, 2, 3, 4 of 5 blokken (de mogelijke hoeveelheden segmenten, gevonden in de tekstjes).





Dit is natuurlijk slechts mijn eigen intuïtie en ervaring. Het wordt pas zinvol als we het kunnen koppelen aan literatuur. Dit doe ik aan de hand van de reeds besproken theorie van Huron (2007). Zie tabel 4 in paragraaf 1.4.6 voor een overzicht met de belangrijkste elementen die hij noemde als het gaat om voorspelbaarheid, verrassing en spanning.

De dingen die ik gehoord heb zijn goed te koppelen aan de theorie van Huron (2007). Ik zie in mijn beschrijving van de stukjes goed de indeling in de drie categorieën voorspelbaarheid, verrassing en spanning terug. De basis van de muziek zie ik als voorspelbaar. Bij elke overgang van blokje naar blokje zit een verrassing die ik kan indelen bij een van de voorbeelden die Huron geeft. Sommige verrassingen zijn “écht” verrassend, door algemene muzikale normen te overschrijden, andere zijn slechts verrassend wat betreft wat zich heeft afgespeeld in deze specifieke muziek. Voorafgaand aan de verrassingen zit in een aantal gevallen een spanningsopbouw, die ook aan Huron gekoppeld kan worden. Bezie tabel 20 voor een herziene versie van mijn beschrijving van de blokjes muziek. In de kolom “Herziene beschrijving” is zichtbaar hoe ik mijn beschrijving heb herschreven, nu enkel in termen van voorspelbaarheid, verrassing en spanning. Zo heb ik bijvoorbeeld bij D de zin “aan het eind een kleine crescendo” vertaald naar “aan het eind een korte spanningsopbouw”. In de laatste kolom staan de specifieke muzikale elementen die ik heb gehoord. Ik heb daarbij in de meeste gevallen letterlijk elementen uit de lijst van Huron (2007) kunnen kiezen. In sommige gevallen hoorde ik echter dingen die niet in mijn schema over Huron voorkomen. Die elementen, die ik dus zelf toegevoegd heb, heb ik in tabel 20 schuingedrukt.

Tabel 20: Analyse van mijn indeling, geïnspireerd op Huron (2007)

	Herziene beschrijving	Specifieke muzikale elementen
A	Voorspelbaar	Metrische hiërarchie Metric proximity Interne herhaling
B	Korte verrassing (stuk-specifiek) Voorspelbaar (algemene normen)	Verandering in instrumentatie Step declination Herkenbaar thema Stabiele harmonische ritmes Ostinaten
C	Verrassend (algemene normen) Verder voorspelbaar (stuk-specifiek)	Onwaarschijnlijke melodische sprong <i>Dissonante toon</i> Interne herhaling
D	Verrassende verandering Aan eind korte spanningsopbouw	Verandering in instrumentatie <i>Crescendo</i>
E	Begint met spanning Na een tijdje voorspelbaar	<i>Suggestief ritme en instrumentatie</i> Interne herhaling
F	Licht verrassende verandering	Wisseling van instrumentatie
G	Grote spanningsopbouw Eindigt met verlossing v.d. spanning	<i>Toenemende dissonantie</i> <i>Crescendo</i> <i>Stijgende lijn in de tonen</i> Stereotype cadens
H	Zeer verrassende verandering Steeds meer spanning toegevoegd	Verandering in instrumentatie Verandering in stijl Verandering in luidheid (zachter) Zeldzame akkoordopeenvolgingen <i>Crescendo</i>
I	Bevredigend loslaten van de spanning Korte verrassing (hh. eerder stukje) Vervolgens voorspelbaar	Stereotype cadens Interne herhaling
J	Zeer verrassend en spannend Uitgebreide verlossing v.d. spanning	Wisseling in instrumentatie Een oplossing uitstellen <i>Crescendo</i> Clichématige cadens

Dat zijn dan ook suggesties voor toevoegingen op de lijst van Huron:

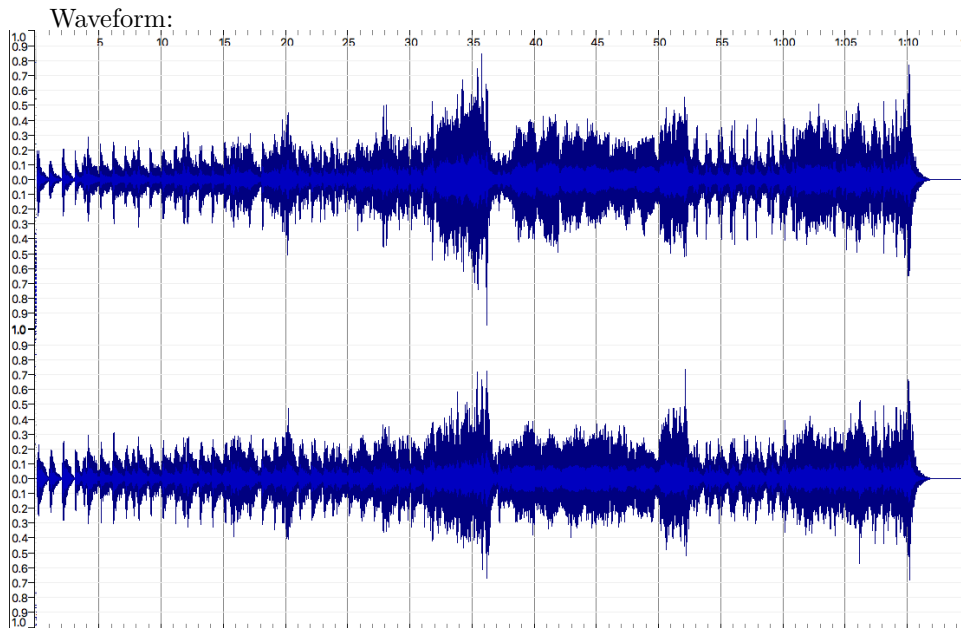
- Dissonante tonen zijn wat mij betreft een zeer duidelijk voorbeeld van spanning.
- Crescendo voelt ook duidelijk als een spanningsopbouw die een verlossing nodig heeft.
- Een stijgende melodische lijn kan ook voelen als steeds intenser wordend.
- Ritme, instrumentatie, klankkleur, melodie kan ook suggestief zijn zonder dat er een specifieke truc dik bovenop ligt. Soms is het lastig om de vinger te leggen op waar het nu precies aan ligt.

3.3 Low-level audio-analyse

Ik ga niet diep in op de low-level audio-analyse. Wel geef ik een aantal visualisaties weer die een goed beeld geven van het fragment en mijn eigen indeling ondersteunen (verkregen met *Sonic Visualizer*). Ik trek geen uitgebreide conclusies uit deze beelden, maar laat dat over aan toekomstig onderzoek.

De eerste weergave die ik geef, is de simpele waveform. Deze geeft de geluidsintensiteit weer. Het nadeel is dat hier weinig informatie uit te halen is. Zo zouden

twee totaal verschillende fragmenten dezelfde waveform kunnen hebben, als ze toevallig dezelfde geluidsintensiteit hebben. Toch kun je er wel een aantal segmenten in terugzien:



Ter vergelijking een herhaling van mijn eigen tijdlijn:



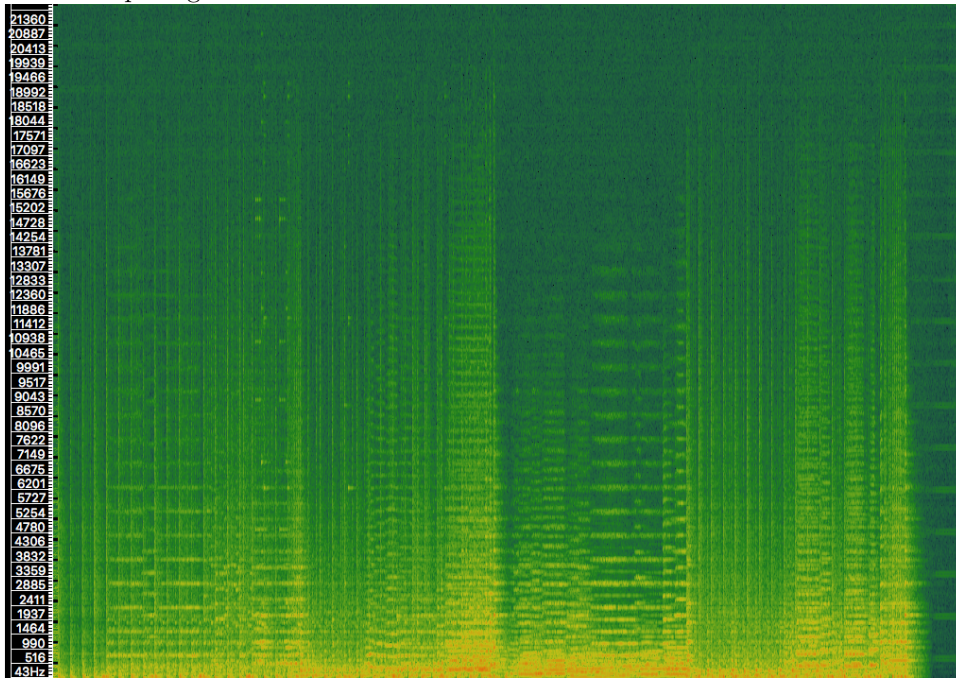
In de waveform zie ik de opbouw naar mijn streep op 0.20 duidelijk terug doordat daar de pieken groter zijn. Hetzelfde geldt voor de piek rond 0.32, met daarna een stukje lage piek tot 0.36. Vervolgens zie ik een redelijk egaal stukje pieken tot 0.50. In mijn opbouw had ik echter de twee seconden erna erbij gerekend. Tot slot zie ik de segmenten van 0.52 tot 1.00 en van 1.00 tot 1.10 ook duidelijk terug.

In het geval van dit specifieke audiofragment is de waveform dus wel degelijk inzichtelijk, maar vooral achteraf gezien. Aan geluidsintensiteit alleen vallen weinig conclusies te trekken.

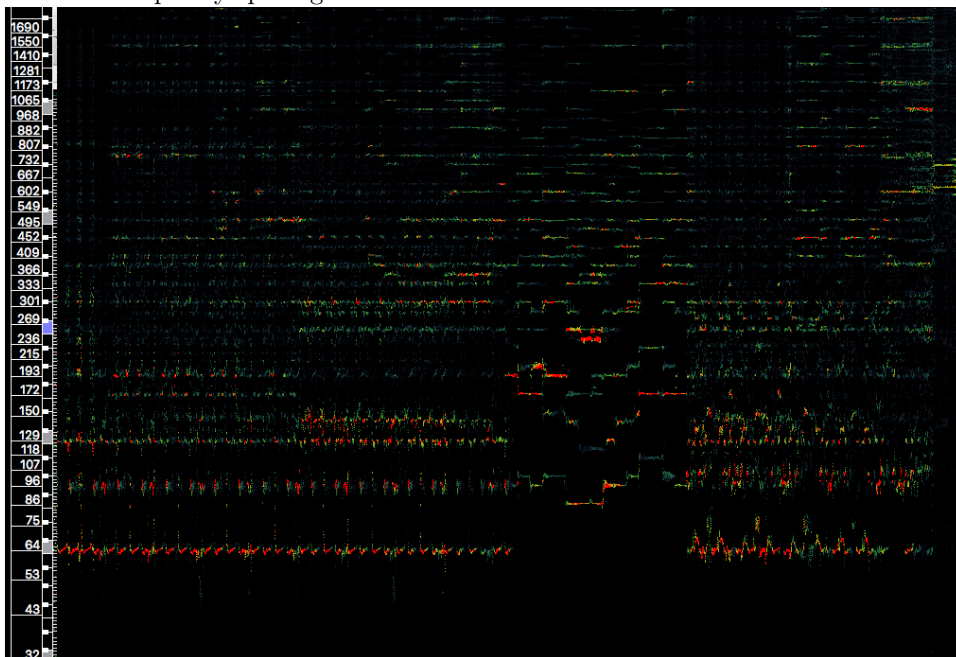
Tot slot toon ik nog twee visualisaties: een standaard spectrogram en een peak frequency spectrogram. Het eerste toont de frequenties tot de helft van het totale bereik van het fragment (hoger zijn het boventonen, die ons gehoor niet of niet bewust hoort). Hier zit nog veel ruis in, waardoor er lastig individuele muzikale kenmerken in te zien zijn. Met segmentatie als doeleinde is er echter wel wat te zien.

De tweede afbeelding is ook een spectrogram maar dan met een filter, waardoor redelijk accuraat de precieze frequenties zichtbaar zijn van individuele tonen. Beide afbeeldingen zijn wederom bedoeld als ondersteuning van de segmentatie.

Plain spectrogram:



Peak Frequency spectrogram:



3.4 Koppeling classificatie - Huron (2007)

Dan wordt het nu tijd om de cinematografische beschrijvingen te koppelen aan de muziek, om de vraag te kunnen beantwoorden die ik mezelf had gesteld: *Op welke manier kan de tekstuele verhaalstructuur verbonden worden aan specifieke momenten in de muziek?*

Hiervoor stel ik een vertaling voor van mijn classificatie van de zinnen naar Huron's onderverdeling van muzikale elementen die spelen met verwachting (en andersom):

Tabel 21: Vergelijking aangepaste Liemclassificatie en Huron (2007)

Klasse	Muzikale termen
1. Geen doel	Voorspelbaar
2. Onduidelijk doel, geen zekerheid	Onvoorspelbaar zonder spanning
3. Duidelijk doel, geen zekerheid	Onvoorspelbaar met spanning
4. Duidelijk doel, zekerheid	Voorspelbaar met spanning
5. Doel niet halen	Verrassing zonder spanningsverlossing
6. Doel wel halen	Verrassing met spanningsverlossing

Nu kan ik een willekeurige beschrijving kiezen en de reeks klassen vertalen naar muzikale termen. Zie bijvoorbeeld nogmaals de tabel met de analyse van beschrijving V:

Tabel 22: User Description V

4	4	4	4	6	1	4	4
---	---	---	---	---	---	---	---

Deze luisteraar zou, volgens mijn vertaaltabel, bijvoorbeeld het volgende gehoord kunnen hebben: een lange spannende episode zonder al te veel onvoorspelbaarheden, toeleidend naar een verlossende verrassing. Vervolgens een kort voorspelbaar stukje, eindigend in weer wat spanning.

Natuurlijk is dit behoorlijk wat speculatie. Bovendien weet je nog steeds niet welke muzikale dingen hij/zij precies gehoord heeft die spannend of voorspelbaar waren. Het zou interessanter zijn om de boel om te draaien: kun je, slechts luisterend/kijkend naar de audio, voorspellen wat luisteraars zouden kunnen horen? Gezien de diversiteit aan klassenreeksen en segmentaties, lijkt dit een lastige opgave.

Met behulp van mijn vertaalsleutel, heb ik mijn eigen analyse van de audio vertaald naar een reeks klassen:

- A 1.
- B 2.
- C 2, 1.
- D 2, 3.
- E 3, 4.
- F 6.
- G 3, 6.
- H 6, 3.
- I 6, 1.
- J 3, 6.

Zo kom ik op de volgende reeks: | 122123346 | 366 | 36 | 136 |, oftewel vier keer een opbouw naar een doel dat behaald wordt. Dat zie ik in geen enkele andere beschrijving terug.

Mijn eigen analyse, op basis van mijn gehoor, in termen van de elementen van Huron (2007), komt dus niet overeen met de verhaalstructuur van de beschrijvingen van de deelnemers. Dit maakt het in dit geval lastig die beschrijvingen aan specifieke momenten in deze muziek te koppelen. Het zegt echter niet dat zo'n verbinding absoluut ondenkbaar is. Er zijn een aantal redenen waarom het in dit geval niet lukt, die in een vervolgonderzoek vermeden kunnen worden.

Ten eerste geeft het onderzoek van Liem et al. (2013) geen garantie dat een deelnemer het stuk van voor naar achter heeft doorgeluisterd. Er werd geen controle uitgevoerd over de luistertijd en er werd zelf als instructie meegegeven om “zo lang als nodig” te luisteren. Bovendien hadden Liem et al. (2013) met hun onderzoek een ander doel; zij vroegen naar een cinematografische scène waar de muziek in zou passen om daarmee narratieve associaties te bestuderen. Ze hebben echter níet expliciet gevraagd om een letterlijke en volledig dekkende transcriptie van de muziek in verhaalvorm.

De luisteraars zouden wellicht veel meer structurele informatie gehoord en beschreven kunnen hebben, als ze een andere opdracht hadden gehad. Voor een antwoord op de vraag “Op welke manier kan de tekstuele verhaalstructuur verbonden worden aan specifieke momenten in de muziek?” zal dus een genuanceerder vervolgonderzoek moeten worden opgezet. Ik zal hiervoor een aantal suggesties doen in hoofdstuk 5.

4 Resultaten

Na vergelijking van de classificaties van Vendler (1967) en Liem et al. (2013), blijken de twee methoden duidelijk niet equivalent te zijn. Het belangrijkste verschil ligt in doelgerichtheid: waar Vendler de focus legt op verschillen tussen predicaten en processen, leggen Liem et al. de focus op verschillen in doelgerichtheid en of de doelen duidelijk gespecificeerd zijn of niet. De classificatie van Liem et al. met de toevoegingen van het wel of niet daadwerkelijk behalen van een doel gaf een duidelijk beeld van de verhaalstructuur, met een redelijk evenwichtige verdeling van klassen.

Door reeksen van klassen te segmenteren op basis van het bereiken van een klasse 5 of 6, is een duidelijk beeld te scheppen van de verhaalstructuur die een luisteraar gehoord heeft. Het analyseren en segmenteren van de tekstjes resulteerde echter in zeer uiteenlopende reeksen van klassen, verschillende segmentaties en geen heersende dominante klasse.

Ook de analyse die ik zelf maakte in termen van muzikale elementen die Huron (2007) noemt, past op geen enkele beschrijving van de deelnemers. Opvallend is dat ik zelf duidelijk een aantal keer een verlossing van de opbouw hoor (er zitten bijvoorbeeld een aantal duidelijke cadensen in het stuk). Mijn structuur bevat dan ook een aantal opbouwen naar een klasse 6. Dit komt echter in geen enkel geval terug in de beschrijvingen van de deelnemers. Dit hoeft echter niet te betekenen dat er überhaupt geen verbinding tussen tekst en muziek mogelijk is, omdat Liem et al. (2013) in hun experiment niet naar een letterlijke transcriptie van de muziek vroegen.

Hoewel in het geval van dit fragment en deze beschrijvingen een verbinding van tekst en muziek lastig is, is er wel een manier om dit soort segmentaties aan bepaalde muzikale kenmerken te koppelen. Daarvoor is het volgende stappenplan toepasbaar: beschrijf eerst de muziek in high-level muzikale kenmerken (zoals bijvoorbeeld die uit tabel 4). Herschrijf die beschrijving vervolgens volledig in termen van voorspelbaarheid, verrassing en spanning. Pas vervolgens de vertaalsleutel uit tabel 21 toe om een reeks klassen te vinden die vergelijkbaar is met de tekstuele segmentaties.

Andersom is het ook mogelijk: neem een segmentatie van een verhalende beschrijving, pas daar de vertaalsleutel uit tabel 21 op toe om een mogelijke beschrijving van de muziek te krijgen in termen van voorspelbaarheid, verrassing en spanning, en bekijk vervolgens welke high-level elementen tot die muzikale structuur zouden kunnen hebben geleid. Daarvoor is tabel 4, met mijn toevoegingen die ik geef op pagina 26, een begin.

De muzikale elementen die dan veel naar voren zouden komen, zijn voor het vakgebied van Music Information Retrieval interessante high-level kenmerken van muziek om op geautomatiseerde wijze uit audiofragmenten te kunnen halen in de toekomst. Een aantal kenmerken die nu al veelbelovend lijken:

- Crescendo
- Stereotiepe cadensen
- Verandering van instrumentatie
- Dissonantie
- Melodische eigenschappen

5 Discussie en suggesties voor vervolgonderzoek

Zoals eerder al aangegeven, is een belangrijke oorzaak voor het feit dat de verhaalstructuur van de teksten niet te verbinden is met de muziek de vraagstelling die Liem et al. (2013) in hun experiment gebruikten. Ze vroegen naar een vrije-vorm beschrijving van een scène die bij de muziek zou passen en gaven de instructie om zo lang als nodig te luisteren. Dit geeft geen garantie dat de deelnemers het hele fragment geluisterd hebben en heeft ervoor gezorgd dat de beschrijvingen geen letterlijke vertalingen van muziek naar beelden zijn. Om een betere verbinding te kunnen maken is dus vervolgonderzoek nodig. Daarvoor is het nuttig eerst nog een kritisch te kijken naar dit onderzoek.

5.1 Discussie

5.1.1 Classificatie

In dit onderzoek heb ik de tekstjes geanalyseerd met de classificatie van Liem et al. (2013), plus twee toegevoegde klassen. Die klassen heb ik uiteindelijk als basis gebruikt voor segmentatie. Maar je zou kunnen zeggen dat klassen 5 en 6 eigenlijk in een andere categorie horen. De klassen 1 t/m 4 zeggen of een bepaalde activiteit een doel heeft, of dat doel duidelijk is en of het duidelijk is dat het doel behaald gaat worden. Klasse 5 en 6 zijn eerder *markeringen* van het wel of niet behalen van een doel. Klasse 5 en 6 kunnen in principe niet voorkomen als er alleen maar klassen 1 (geen doel) aan vooraf gaan. Dat komt inderdaad niet voor in de beschrijvingen in dit onderzoek. Bovendien is het niet zeker of een zin in een tekst altijd alleen als een 5 of een 6 geanalyseerd moet worden; misschien kan een zin van klasse 4 ook een 6 impliceren of andersom? Denk bijvoorbeeld aan de zin “Hij rende zijn huis in om te ontsnappen aan zijn achtervolgers”. Het lijkt lastig om te kiezen tussen klasse 4 (doelgericht met zicht op behalen van het doel) en klasse 6 (doel behalen).

Een idee voor vervolgonderzoek zou zijn om de categorieën te scheiden en aan elke zin zowel een klasse 1, 2, 3 of 4 als een klasse 5 of 6 te hangen. Misschien handig om klasse 5 en 6 dan A en B te noemen. Echter, een zin met klasse 1 kan niet ook een A of B zijn; het is immers een doelloze zin, dan is er geen doel om (niet) te behalen. In dat soort gevallen nog een derde klasse O toevoegen is mogelijk, maar maakt het wel weer wat omslachtig.

Een beter idee is om tóch te kiezen voor alleen de originele klassen 1 t/m 4 van Liem et al. (2013). Het daadwerkelijk falen of slagen in het behalen van het doel kan dan dienen als markering; het zit dan niet in de classificatie maar is een extra element van de analyse.

5.1.2 Chronologie en klasse 1

Een ander probleem waar ik in dit onderzoek tegenaan liep, was het probleem van chronologie. Het was door de vrije opdracht niet bij alle beschrijvingen duidelijk of de beschreven beelden op volgorde van de tijd beschreven werden en of de muziek er noodzakelijk mee gelijkliep. Dat is in vervolgonderzoek makkelijk op te lossen door gerichtere vragen te stellen.

Maar, hieraan gerelateerd, sommige zinnen zeggen iets over context, en niet over de actie van de beelden. Het verwisselen van die zinnen met andere zinnen, zal niets aan de betekenis van het geheel veranderen. Het lijkt alsof dat soort zinnen altijd van klasse 1 zijn. Volgens mij is er dus een tweedeling binnen klasse 1: aan de ene kant zinnen die echt een state uitdrukken, waarbij de bijbehorende beelden wel belangrijk maar niet doelmatig zijn en waarschijnlijk muziek hoort die geen spanning opbouwt, zonder verrassingen. Aan de andere kant zijn er de “contextzinnen”, die alleen dienen ter verduidelijking, niet voorkomen in de beelden en verwisselbaar

zijn met andere zinnen zonder de betekenis te veranderen. De contextzinnen zullen geen parallel hebben in de muziek, dus is het vervolgonderzoek wellicht handig om ze niet mee te nemen in de uiteindelijke analyse.

5.1.3 Segmentatie

Met het wel of niet behalen van een doel als markering, is een goede, inzichtelijke segmentatie te maken. Maar is dit de enige mogelijkheid voor segmentatie? Bij beschrijving II (tabel II en bijlage A en B) kwam ik een aantal zinnen tegen die ook een segmentgrens suggereerden, zonder een doel wel of niet te behalen. Een voorbeeld van zo'n zin: "When the music breaks at 30 seconds in, a cut to his parents fighting." Dit is duidelijk zowel een scène- als een muzikale overgang. In vervolgonderzoek moet hier rekening mee gehouden worden (in de vraagstelling, of eventueel in de vorm van een extra markeringsteken voor dit soort segmentgrenzen).

5.1.4 Verbinding tekst en muziek

De high-level muzikale kenmerken die Huron (2007) noemt, zijn een goed begin voor een lijst met elementen die mogelijk muzikale verhaalstructuur tot stand brengen. Met mijn gehoor heb ik er echter nog een aantal dingen bijgehaald. Mogelijk zijn er bij andere fragmenten nog andere dingen te horen. De lijst in tabel 4 is dus geen vaststaande lijst, maar slechts een nader aan te vullen begin. De driedeling van voorspelbaarheid, verrassing en spanning is wel goed aan verhaalstructuur te verbinden (zie bijvoorbeeld mijn vertaalsleutel in tabel 21).

Liem et al. (2013) benoemen in hun artikel het feit dat deelnemers in hun keuze voor een bepaalde scène ook stappen zetten die niet aan de muziek zelf gerelateerd zijn (maar bijvoorbeeld aan een herinnering). Dit zal betekenen dat een volledige verbinding van tekst en muziek niet mogelijk is; een fragment heeft niet één enkele ware verhaalstructuur waar slechts een eindige lijst met beelden op past. Muziek blijft tot de verbeelding spreken en zal altijd voor elk individu iets anders betekenen.

5.2 Suggesties voor vervolgonderzoek

5.2.1 Vraagstelling

De vraagstelling naar de deelnemers toe is de belangrijkste sleutel naar een onderzoek met meer mogelijkheden. Ten eerste moeten de deelnemers het hele fragment, liefst meerdere malen, beluisteren. Wat ze vervolgens beschrijven van hun scène moet een één-op-één relatie hebben met de muziek. Hoe bereik je dat, zonder de deelnemers te veel te sturen en het verband tussen tekst en muziek eigenlijk zelf van te voren al te creëren? Is het voldoende om simpelweg zeggen "Luister dit fragment drie keer en beschrijf een cinematografische scène die hier volledig bij past"? Of hebben luisteraars meer sturing nodig?

Ik zou voorstellen om de deelnemers eerst één keer het fragment in zijn geheel te laten luisteren. Vervolgens vraag je ze nog een keer luisteren en daarbij aan te tekenen hoeveel delen ze horen in de muziek. Dan kunnen ze gaan nadenken over filmbeelden waar de muziek goed bij zouden passen. Zodra ze een goed beeld hebben van het verhaal, moeten ze nog een laatste keer naar het fragment luisteren voor ze het opschrijven.

Hiermee vraag je wel een flinke tijdsinvestering van de deelnemers. Het is echter wel nodig dat ze het fragment een aantal keer gehoord hebben om een goede verbinding te kunnen maken met wat ze horen en wat ze opschrijven. Ik denk dat deze methode een goede middenweg is tussen het vrijlaten en het sturen van de deelnemers.

Achteraf kun je de deelnemers ook nog vragen wat hen precies triggerde voor een bepaalde verdeling te kiezen. Het is interessant om te zien wat ze noemen als je ze vraagt naar muzikale eigenschappen die tot een segmentgrens leidden. Waarschijnlijk kunnen ze het wel beschrijven, ook al hebben ze geen muziektheoretische kennis.

5.2.2 Tekstanalyse

Voor het analyseren van de teksten, stel ik de volgende classificatie voor:

- 1 Activiteit of staat zonder doel
 - i Deel van het verhaal
 - ii Contextbeschrijving
- 2 Activiteit met onduidelijk doel, zonder zekerheid van het behalen van het doel
- 3 Activiteit met duidelijk doel, zonder zekerheid van het behalen van het doel
- 4 Activiteit met duidelijk doel, met zekerheid van het behalen van het doel

Naast deze classificatie kan gebruik gemaakt worden van markeringen:

- . Het niet behalen van een doel
- ! Het wel behalen van een doel
- ; Een andere vorm van segmentatie die naar voren komt in de tekst

5.2.3 Vertaling naar muzikale verwachting

Als met behulp van de classificatie en markeringen reeksen geformuleerd zijn die de verhaalstructuur weergeven, kunnen die reeksen vertaald worden in termen van muzikale verwachting. Dit kan bijvoorbeeld op dezelfde manier als ik in tabel 21 doe. Daarmee kan bepaald worden of dit verbonden kan worden aan de muziek, door onder andere de lijst van Huron (2007) te gebruiken. Eventueel zouden die resultaten vergeleken kunnen worden met eigen redenen van deelnemers voor hun segmentatie.

5.2.4 Low-level vervolg

Als daar opvallende elementen uit naar voren komen, kan ook een verbinding gezocht worden met de low-level kenmerken uit het audiosignaal. Dat kan door gevonden high-level elementen uit audiosignaal te extraheren. Andersom is ook mogelijk: neem een fragment, kijk naar low-level kenmerken, maak daar een muzikale verhaalstructuur van in termen van voorspelbaarheid, verrassing en spanning, verzin daar een verhaal bij (zonder het fragment zelf te luisteren) en kijk of mensen dat aan het juiste fragment kunnen verbinden.

6 Samenvattende conclusie

Deelvraag 1. Wat is de positie van dit onderzoek binnen het onderzoeksveld?

Dit onderzoek bevindt zich in het zeer interdisciplinaire onderzoeksveld van *Music Information Retrieval* (MIR). Casey et al. (2008) maken onderscheid tussen methoden gebaseerd op metadata en methoden gebaseerd op muzikale inhoud, waarbij bovendien onderscheid wordt gemaakt tussen high-level en low-level kenmerken. De moeilijkheid van Music Information Retrieval zit in het overbruggen van de kloof tussen low-level en high-level kenmerken, tussen audiosignaal en muziektheoretische concepten. In dit onderzoek analyseerde ik metadata en beschreef de high-level kenmerken van de bijbehorende muziek. In de toekomst kan de verbinding gemaakt worden met de low-level kenmerken.

Deelvraag 2. Wat zijn de mogelijkheden voor een tekstuele analyse van vrije-vorm beschrijvingen van verhalende beelden?

Ik heb twee classificaties besproken, waarmee cinematografische beschrijvingen geanalyseerd kunnen worden. Ten eerste de classificatie van Vendler (1967):

- A State: predicaat dat waar kan zijn gedurende een bepaalde tijd.
- B Achievement: predicaat dat waar kan zijn op een bepaald moment.
- C Activity: proces, waarbij elk deel van dezelfde aard is als het geheel.
- D Accomplishment: proces met een bepaald eind.

Ten tweede de classificatie van Liem, Larson & Hanjalic (2013):

- 1 Activity or state with no goal
- 2 Activity with underspecified goal and no certainty of achieving the goal.
- 3 Activity with well-specified goal and no certainty of achieving the goal.
- 4 Activity with well-specified goal and certainty of achieving the goal.

Het verschil ligt in de focus op doelmatigheid. Zelf heb ik aan de laatste classificatie nog twee klassen toegevoegd:

- 5 Not achieving the goal (whether as a surprise or not)
- 6 Achieving the goal (whether as a surprise or not)

Hierdoor zijn segmentaties te maken, door een tekst zin voor zin te analyseren en de gevonden reeks van klassen op te delen in segmenten die naar een 5 of een 6 toewerken.

Deelvraag 3. Wat zegt literatuur over muzikale verwachting?

Muziek speelt met de verwachtingen van de luisteraar en wekt daarmee allerlei reacties en gevoelens op. Huron (2007) noemt drie elementen die meespelen in dit spel met verwachting: voorspelbaarheid, verrassing en spanning. Voorspelbaarheid kan leiden tot het *prediction effect*: de luisteraar ervaart plezier doordat zijn verwachtingen uitkomen. Het fenomeen *contrastive valence*, een positieve reactie onmiddellijk volgend na een negatieve reactie, maakt verrassingen ook plezierig. Door te spelen met voorspelbaarheid en verwachting, kan een gevoel van anticipatie bereikt worden en spanning worden opgebouwd. Voorspelbaarheid, verrassing en spanning kan op vele manieren bereikt worden, door te spelen met schematische (algemene of culturele) normen of dynamische normen (binnen het stuk zelf vastgesteld), door te spelen met wat vertrouwd is bij de luisteraar of met verwachtingen waar de luisteraar zich bewust van is. In tabel 4 op pagina 14 is een overzicht te

vinden van wat Huron (2007) specifiek voor muzikale voorbeelden noemt.

Deelvraag 4. Is er een “gemiddelde” onderliggende verhaalstructuur in de cinematografische beschrijvingen bij een muziekstuk?

Nadat ik tekstuele beschrijvingen zin voor zin een klasse heb toegewezen en segmentatie heb toegepast na elke klasse 5 of 6, ben ik tot 11 zeer verschillende reeksen van klassen gekomen. Ik heb gezocht naar overeenkomsten tussen de reeksen, door te kijken naar de klassenopeenvolgingen zelf, naar de segmentatie en naar dominante klassen. Alledrie die methodes hebben niet geleid tot een duidelijk onderliggend verband. Dit kan echter zeer goed komen door de manier waarop de deelnemers gevraagd werden hun beschrijving te geven.

Deelvraag 5. Op welke manier kan de tekstuele verhaalstructuur verbonden worden aan specifieke elementen in de muziek?

Door de geanalyseerde reeksen te vertalen in termen van voorspelbaarheid, verrassing en spanning (zie tabel 21 op pagina 29), kan de tekstuele verhaalstructuur aan muzikale opbouw van verwachting gekoppeld worden. Vervolgens kan in de muziek gezocht worden naar high-level kenmerken die die muzikale opbouw tot stand gebracht zouden kunnen hebben. Met deze beschrijvingen in dit onderzoek was de verbinding echter lastig te maken, omdat de deelnemers de scènes niet hadden geschreven om letterlijk chronologisch met de muziek mee te lopen.

Hoofdvraag: Wat zijn veelbelovende resultaten uit mijn onderzoek, die meegenomen kunnen worden in vervolgonderzoek?

Een van de belangrijkste ontdekkingen van dit onderzoek, is dat de classificaties van Liem et al. (2013) en Vendler (1967) absoluut niet equivalent zijn. De doelgerichtheid maakt het verschil, waardoor de classificatie van Liem et al. een vergelijking met muziek mogelijk maakt. Ik heb zelf die classificatie uitgebreid met twee klassen, waardoor de verdeling van klassen veel evenwichtiger wordt. Voor vervolgonderzoek is het echter wellicht handig om het te houden bij de originele 4 klassen. Er zouden dan wel markeringen toegevoegd moeten worden tijdens de analyse, die aangeven wanneer een er een doel wel of niet behaald wordt.

Een ander resultaat is het succes van de theorie van Huron (2007). Zijn beschrijving van muzikale verwachting in termen van voorspelbaarheid, verrassing en spanning maakt een vergelijking tussen tekstuele verhaalstructuur en muzikale structuur mogelijk.

Verder heb ik door dit onderzoek waardevolle suggesties voor vervolgonderzoek opgedaan. Deze zijn te vinden in hoofdstuk 5.

Referenties

- [1] Casey, M. A., Veltkamp, R., Goto, M., Leman, M., Rhodes, C. & Slaney, M. (2008). Content-Based Music Information Retrieval: Current Directions and Future Challenges. *Proceedings of the IEEE* 96(4), 668-696. doi: 10.1109/JPROC.2008.916370
- [2] Champollion, L. & Krifka, M. (2015). Mereology. In P. Dekker & M. Aloni (eds), *Cambridge Handbook of Semantics* (pp. 512-541). Retrieved from <http://ling.auf.net/lingbuzz/002099>
- [3] Huron, D. (2007). *Sweet Anticipation: Music and the Psychology of Expectation* (First MIT Press paperback edition). Cambridge, Massachusetts: The MIT Press.
- [4] Liem, C. C. S., Larson, M. & Hanjalic, A. (2013). When music makes a scene: Characterizing music in multimedia contexts via user scene descriptions. *Int J Multimed Info Retr* 2(2013), 15-30. doi:10.1007/s13735-012-0031-3
- [5] Maienborn, C. (2010). Event semantics. In C. Maienborn, K. von Stechow & P. Portner (eds), *Semantics. An international handbook of natural language meaning* (pp. 1-27). Retrieved from http://www.uni-tuebingen.de/fileadmin/Uni_Tuebingen/SFB/SFB_833/A_Bereich/A1/Maienborn-2010-HSK_Event_semantics.pdf
- [6] Vendler, Z. (1967). Verbs and Times. *The Philosophical Review* 66(2), 143-160. Retrieved from <http://links.jstor.org/sici?sici=0031-8108%281967%29%3A2%3C143%3AVAT%3E2.0.CO%3B2-2>