

Metoder for å analysere video-opptak av musikk og bevegelse

Egil Haga

Innledning

I denne teksten skal jeg presentere en tilnærming til å analysere musikk og bevegelse. Tilnærmingen innebærer på den ene side utvikling av et sett med deskriptive termer, og på den andre side forslag til software som man kan bruke som verktøy i analysen. Det skal bemerkes at tilnærmingen er utarbeidet som en metode for å analysere musikk og bevegelse slik de framstår på et (digitalt) video-opptak.

Framstillingen har utgangspunkt i et pågående doktorgradsarbeid der jeg analyserer video-opptak av dansere som beveger seg til musikk. Hensikten med analysene er å drøfte spørsmål knyttet til musikk-bevegelse korrespondanser innenfor relativt korte tidsrom (15-20 sekunder): *"Når vi sier at musikk og bevegelse 'ligner' på hverandre, hvilke trekk i musikken og bevegelsen sikter vi til?"*

En diskusjon av korrespondanse mellom musikk og bevegelse kan åpenbart angripes fra flere vinkler. Jeg har valgt å begrense fokuset til såkalte ikke-symbolske aspekter ved musikk og bevegelse, slik som dynamikk-kinematikk, rytme/frasering, synkroniseringspunkter, "chunking", "effort" og aktivering. I denne artikkelen skal jeg primært konsentrere meg om å gjøre rede for hva som menes med "chunking" og aktivering og hvordan disse termene anvendes i en analysetilnærming.

Analysetilnærmingen er i utgangspunktet utviklet for å kunne beskrive likheter og ulikheter i samtidige musikalske og bevegelsesmessige forløp. Dette mener jeg bør være svært relevant i ulike musikkterapisammenhenger, der musikk og bevegelse opptrer på samme arena og skaper et felles opplevelsrom. Innenfor dette rommet, for eksempel innenfor rammen av terapeutisk improvisasjon, synes korrespondanser, eller imitasjon, mellom musikalske og gestiske innspill og utspill å være et sentralt omdreiningspunkt. Det er i denne forbindelse nærliggende å nevne den rollen sensitivitet til kryss-modal matching tillegges av Daniel Stern i forbindelse med *affektinntoning* (Stern 2000), og videre hvordan *affektinntoning* diskuteres som en vesentlig komponent i en terapeutisk prosess (Trondalen & Skårderud 2007).

Den åpenbare utfordringen med å beskrive musikk og bevegelse er at begge er fenomener som karakteriseres ved måten de utfolder seg i tid. "Problemet" synes todelt:

- Tidsavhengige fenomen framstår som ”flyktige”; det er som om de blir borte for oss på samme tid som de dukket opp, og
- de karakteriseres av forandringer i tid. Dette er endringer vi er svært sensitive til. Et eksempel kan være en ørliten klangforandring i musikk eller en umerkelig ’nøling’ i en bevegelse, og som vi vet er helt essensielle for opplevelsen vår, men som vi normalt ikke er i stand til å fange inn med nøyaktige beskrivelser.

I det følgende skal jeg foreslå hvordan dette ”problemet” kan håndteres, om ikke løses, i analyse. Det første grepet som gjøres er å bryte opp forløpet i kortere tidsbiter (”chunks”) for deretter på forskjellige måter å karakterisere endring innenfor disse avgrensede, oppklippede tidsvinduene.

Chunking og endringskarakteristikker

Chunking refererer til det aspektet ved en perseptuell prosess der vi oppfatter et forløp i tid som avgrensede enheter, der hver enhet har et start- og et endepunkt. Dette gjør at vi for eksempel oppfatter en tonerekke som en avgrenset melodilinje som begynner et sted og som stopper et sted, og som så avløses av en ny melodilinje. Dvs. enkelttoner bindes sammen til en helhet.

Tilsvarende oppfatter vi en strøm av bevegelser i avgrensede bevegelse-enheter. Still-serien i figur 1 illustrerer en armbevegelse der danseren løfter armen ut ifra kroppen og fører den tilbake igjen. Denne langsomme ’bølge’-bevegelsen gjentas. Hver av disse ’bølgene’ oppfattes som én ”chunk”.

Bølge 1



Bølge 2

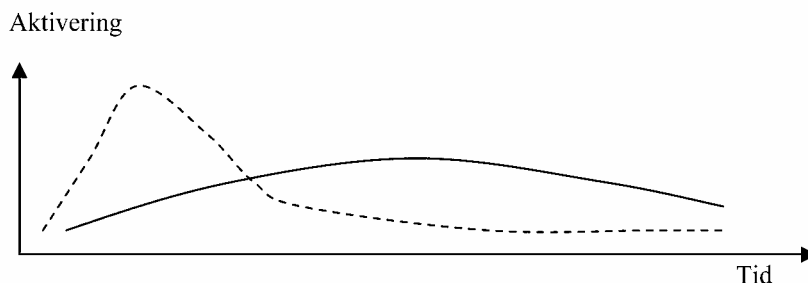


Figur 1. To bølgebevegelser.

Chunking innebærer altså en måte å få grep om flyktige fenomener. Alternative termer som har vært brukt innenfor musikkforskning for denne delen av den perseptuelle prosessen er *segmentering* eller *gruppering* (Bregman 1990; Deutsch 1999). Man har blant annet vært opptatt av å studere hvilke egenskaper i en tonerekke som gjøre at enkelttoner bindes sammen slik at de framstår som en melodi.

Begrepet *chunking* har opphav i den typen strategi vi bruker for å huske lange tallrekker, for eksempel et telefonnummer, ved å binde enkelttallene sammen til meningsfulle enheter (1 og 2 blir til 12), slik at det totalt blir færre enheter å bry hukommelsen med (Snyder 2000). *Chunking* innebærer derfor et element av mening, dvs. vi tillegger eller påtvinger den kontinuerlige strømmen av inntrykk en forståelse, som gjør at den brytes ned til mindre enheter. Av denne grunn foretrekker jeg begrepet fordi det bærer i seg en forståelse av at det vi kunne assosiere med orden/struktur (avgrensing i tid) er nært forbundet med – og avhengig av – en form for meningsinnhold. Med andre ord, å bryte ned en kontinuerlig strøm i mindre biter er vevd inn i det å oppfatte mening.

Den formen for mening jeg sikter til her, og som jeg skal kommentere videre, kan forstås som *ikke-symbolsk* (hva som menes med *ikke-symbolsk*, se under). Jeg skal konsentrere meg om oppfattelsen av *aktiveringskonturer* i musikk og bevegelse, dvs. hvordan chunking av en kontinuerlig strøm av hendelser, musikalske eller bevegelsesmessige, har grunnlag i hvordan vi oppfatter konturer av aktivering i denne strømmen. Termen *aktiveringskontur* er introdusert av Daniel Stern i hans teorier og undersøkelser av tidlig kommunikasjon mellom foreldre og spedbarn (Stern 2000).¹ Aktiveringskontur refererer til hvordan energi, intensitet eller kraft i en bevegelse (eller et lydlig uttrykk) endrer seg innenfor et tidsvindu. Man kan f.eks. tenke seg at bevegelsen settes i gang med en kraftig impuls og at intensitet avtar gradvis etter hvert (stiplet linje), eller at intensitetsforandringen har mer karakter av en bue-kontur (heltrukken linje), som forsøkt antydnet i figur 2:



Figur 2. To ulike aktiveringskonturer.

¹ Første gang publisert i 1985.

Videre er aktiveringskontur forstått som en egenskap ved *vitalitetsaffekter*. Begrepet vitalitetsaffekt innebærer en anerkjennelse av at følelseskategorier som glede, sinne, behag ikke er uttømmende for hvordan vi oppfatter hverandre. Det er tusen måter å være glad på og det er minst like mange måter å være sint på, og svært ofte rekker vi ikke å tenke på om vi (eller den andre) var glad eller sint, men responderer direkte på hvordan gleden eller sinnet ble uttrykt. Man kan f.eks. tenke seg et uendelig antall varianter av hvordan et smil sprer seg ut over ansiktet, fra et ”eksplosivt”, plutselig smil, til et smil hvor smilet sprer seg fra munn til øyne mer gradvis. Fra dagliglivet kjenner vi også til alle måtene en dør kan lukkes på og hvordan slamring vs. et forsiktig, nølende skyv uttrykker nyanser i intensjoner.

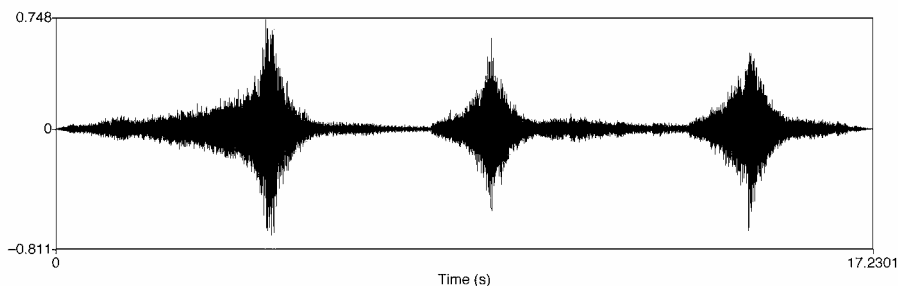
Her skal jeg kort kommentere distinksjonen mellom *symbolsk* og *ikke-symbolsk* mening. Et symbol forstås som et tegn (et snakket eller skrevet ord, en gest, et bilde) som står istedenfor noe annet (f.eks. ordet ’bil’ som refererer til en fysisk gjenstand) der symbolet primært får sin mening gjennom konvensjon (og ikke ved tegnets form). *Ikke-symbolsk* mening er knyttet til de nyansene jeg har beskrevet over (f.eks. nyanser i et smil) og får mening i måten vi responderer på disse nyansene (i handling eller i fornemmelsen av nyanser). For det førspråklige barnet som lever i en verden uten ord, og ennå ikke har navn på tingene og opplevelsene, vil det være nærliggende å omtale det meningsfulle i samspillet med menneskene rundt som *pre-symbolsk* (eller *pre-verbalt*). Men for oss som bruker språket vil *ikke-symbolsk* være en bedre term. Her er det verdt å merke seg at distinksjonen *symbolsk* vs. *ikke-symbolsk* må forstås teoretisk, dvs. i det virkelige liv vil *ikke-symbolske* og *symbolske* meningsaspekter være infiltrert i hverandre – og ”smitte over” på hverandre.

Sterns poeng er at vitalitetsaffekter og hvordan disse karakteriseres av ulike aktiveringskonturer er en viktig bestanddel i den preverbale kommunikasjonen mellom foreldre og spedbarn, som ikke ennå har tilegnet seg språk. Foreldre og barn kommuniserer ikke gjennom symboliseringer, men gjennom å respondere på hvordan bevegelser eller lyder (f.eks. vokale) er karakterisert ved aktiveringskonturer. En viktig forutsetning for kommunikasjonen er at både foreldre og barn er sensitive til nyanser i sine egne og andres utspill. En annen viktig side ved kommunikasjonen er at det viser seg at den foregår på tvers av sansemodaliteter, dvs. at en bevegelse hos barnet blir ”kommentert” med en lyd av mor/far. En forutsetning for dette er at barnet er i stand til å oppfatte oversettelse fra en sansemodalitet til en annen, at bevegelsen (det jeg ser/kjenner at jeg gjør) ”ligner” på lyden (det jeg hører). Stern viser i denne forbindelse til spedbarnsforskning som tyder på at evnen til å integrere inntrykk til ulike modaliteter er medfødt, eller i det minste noe som etableres på et svært tidlig utviklingsstadium.²

² Stern viser til forskning som var tilgjengelig i 1985 da *The Interpersonal World of the Infant* først ble utgitt (Lewkowicz & Turkewitz 1980; Meltzoff & Borton 1979). Kunnskapen om sanseintegrasjon hos spedbarn er utdypet i nyere forskning (Lewkowicz & Kraebel 2004).

På bakgrunn av dette forstår jeg aktiveringskontur som en endring av intensitet eller energi innenfor et tidsvindu, som karakteriserer en bevegelseschunk (både den jeg ser hos andre, og den jeg kjenner at jeg gjør med min egen kropp), eller et avgrenset lydforløp (som jeg hører, eller som jeg frambringer selv, musikalsk lyd inkludert).

I forbindelse med musikalsk lyd vil aktiveringskontur f.eks. referere til hvordan intensitet endrer seg innenfor en frase. I noen musikkformer blir en melodilinje typisk spilt med en økning i intensitet fram mot litt ”etter midten” av frasen og med en påfølgende, ”lukkende” avfrasering (avtagende intensitet), som i en crescendo-decrescendo kontur. I mindre tradisjonelle genre kan slike konturer utgjøre det bærende musikalske elementet, f.eks. som i Varese’s *Amerique*, her illustrert ved en bølgeform-representasjon fra et parti der orkesteret framfører en serie av påfølgende ’utbrudd’:



Figur 3. Bølgeform-representasjon av utdrag fra Varese’s *Amerique*.

Her er det verdt å kommentere aspekter ved slike forløp som er lite (eller overhodet ikke?) kommentert av Stern i forbindelse med aktiveringskonturer. Det synes som om det i utførelsen og oppfattelsen av slike konturer er innebygd en forståelse av en form for intensjonalitet; det er som om vi ”er på vei mot noe”, at vi ”når fram til noe” og at vi ”kommer tilbake/trekker oss tilbake”.³ Det kan synes som om forløpet går gjennom faser, jeg skal bruke termene *prefiks – mål-punkt – suffiks*.⁴

Tilsvarende terminologi finner vi i forskning på non-verbal kommunikasjon. Her forstår man en gest som en kroppsbevegelse som er avgrenset (mht. et definert start- og slutt punkt i tid) og som oppfattes som velformet. Med velformet mener man at bevegelse oppfattes som en ”ekskursjon” fra et startpunkt til et slutt punkt (Kendon 2004). Bevegelsen i figur 1 har et slikt forløp, dvs. armen løfter seg ut fra kroppen og vender tilbake til utgangspunktet. Innenfor dette forskningsfeltet bruker man termene *preparation – stroke – retraction* for å beskrive de fasene som en velformet gest går gjennom.

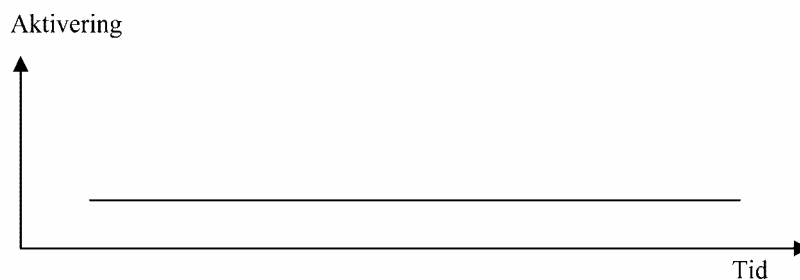
³ Se Snyder (2000:62-63) om *closure* i musikalske fraser.

⁴ Termene er foreslått av Rolf Inge Godøy i samtale.

Uavhengig av hvilke termer man velger, vil jeg foreslå at dette faseaspektet, sammen med en overordnet aktiveringskontur, bidrar til oppfattelse av *chunks* i bevegelse og musikk, og på samme tid utgjør et viktig meningsaspekt i ikke-symbolisk forstand.

Hos Stern blir den pre-verbale ytringens aktiveringskontur omtalt som *dynamisk profil*. Det kan synes som om Stern i første rekke er opptatt av det *dynamiske* aspektet ved slike ytringer. I forbindelse med menneskelig bevegelse er *dynamikk* forstått som de krefter som forårsaker/setter i gang og påvirker/rammer inn bevegelser innenfor et tidsrom (Rosenbaum 1990; Winter 2005). *Dynamikk* begrepet opptrer gjerne i par med bevegelsens *kinematikk*, som er forstått som de aspekter ved bevegelse som er direkte observerbare (banen i rommet, retnings- og hastighetsforandringer). *Dynamikk-kinematikk* distinksjonen har opphav i fysikk/mekanikk, da med referanse til ikke-levende objekters bevegelse, men er videreført inn i studier av biomekanikk (Winter 2005), motor kontroll (Rosenbaum 1990) og visuell persepsjon av kroppsbevegelse (Runeson & Frykholm 1983).

Jeg nevner dette fordi det er ikke alltid slik at det er aktiveringskontur, forstått som et dynamisk aspekt ved en ytring, som tjener til å avgrense/karakterisere en chunk. Bevegelsen som er illustrert med stillserien i figur 8 er et eksempel på dette. Bevegelsen, de to 'bølgene', er langsom, glidende og utført med utholdt, 'jevnt' tilført kraft. Aktiveringskonturen ville kunne illustreres med en slik tegning:



Figur 4. Aktiveringskontur ved utholdt kraft.

Jeg vil foreslå at det i dette tilfellet er det kinematiske aspektet ved bevegelsen, dvs. banen i rommet og måten den blir gjentatt, som er avgjørende for hvordan vi avgrenser bevegelsen. Dette aspektet ved bevegelsen vil jeg omtale som *kinematisk kontur*, og samtidig foreslå at en bevegelse i ulik grad kan karakteriseres ved *dynamiske* og/eller *kinematiske konturer*.

Begrepsparet er også forsøkt anvendt i musikkteori der tempoendringer og tonehøydeendringer forstås som *kinematiske* aspekter, mens *dynamikken* er de underliggende "krefter" som forårsaker disse endringene (Feldman, Epstein & Richards 1992; Todd 1995). Denne anvendelsen kan være problematisk fordi den aktiverer en mer prinsipiell diskusjon om forholdet mellom musikk og

bevegelse; dvs. ved å overføre disse begrepene direkte på musikk introduserer vi forestillinger om at musikken beveger seg i bane i rommet (kinematikk) og at musikken har masse som krefter virker inn på (dynamikk). Jeg skal ikke gå videre inn i en slik diskusjon, men viser til Eric Clarkes bok *Ways of Listening* der temaet er drøftet (Clarke 2005).

Tross disse innvendingene vil jeg bruke begrepene slik det er antydnet over, altså:

- ved at kinematisk kontur i musikk refererer til tonehøydeendringer og tempoendringer,
- og ved at dynamisk kontur i musikk refererer til endring i intensitet/styrke.

Eksemplifisert vil en enkel kinematisk kontur fortone seg slik:



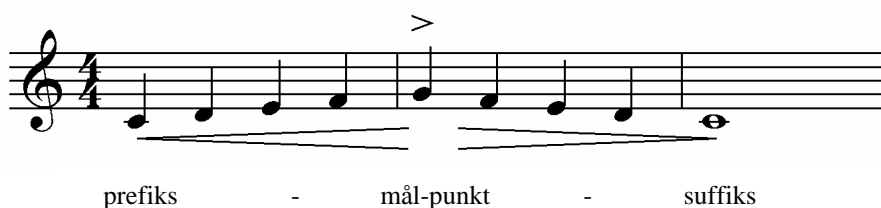
Figur 5. Melodilinje som kinematisk kontur.

mens en dynamisk kontur høres i et slikt eksempel:



Figur 6. Crescendo-decrescendo som dynamisk kontur.

Svært ofte vil flere endringer skje på samme tid slik at forløpet både tilbyr en dynamisk og en kinematisk kontur, i tillegg til at det framstår med en fase-karakteristikk (som diskutert tidligere):



Figur 7. Melodilinje og crescendo-decrescendo som kombinasjon av kinematisk og dynamisk kontur

Chunking – verktøy for å bryte opp auditive og visuelle forløp

I denne delen skal jeg foreslå hvordan man kan gå fram for å bryte opp forløp av musikk og bevegelse i kortere chunks. Det spørsmålet vi stiller oss når vi observerer bevegelsene (i et video-opptak) er: *Hvor oppfatter jeg at en bevegelse begynner, hvor slutter den, og hvor begynner en ny bevegelse?* Tilsvarende spørsmål for musikkforløpet er: *Hvor oppfatter jeg at en sammenhengende musikk-chunk begynner, hvor slutter den, og hvor begynner en ny?* Alternativt kan man spørre seg: *Hvilke elementer hører sammen og er deler i en helhet?*

I første del av teksten diskuterte jeg sammenhengen mellom chunking og endringskarakteristikker. I lys av dette vil jeg foreslå at vi under observasjon av musikk/bevegelse avgjør start og slutt-punkter til chunks på grunnlag av følgende trekk:

- Dynamiske (f.eks. aktiveringskontur) og kinematiske konturer
- Fase-forløp (peak structure)

Oppfattelsen av chunks er i figur 8 illustrert, eller ”frosset”, ved å bruke en still-serie.



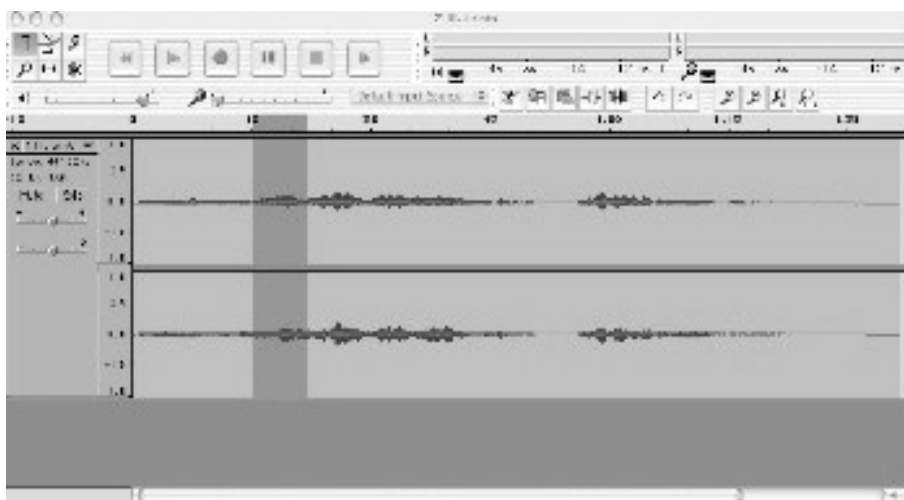
Figur 8. To påfølgende chunks

En slik still-serie kan på grunnlag av et video-opptak lages ved å bruke Quick Time Pro (heretter QTP).⁵ Et annet alternativ er å klippe opp den kontinuerlig video-strømmen i mindre biter, en oppklipping som svar på spørsmålet ”hvilke bevegelsessekvenser hører sammen til én enkelt gest, avgrenset fra andre gester?” Disse oppklippede video-snuttene kan senere brukes til å studere, og karakterisere, en og en bevegelse om igjen og om igjen. Dette skal jeg komme tilbake til.

⁵ Quick Time er et gratis program (tilsvarende Windows Media Player for Windows). Oppgraderingen til Quick Time Pro som muliggjør enkel redigering av lyd og bilde, gjøres ved å kjøpe lisens for et mindre beløp, se www.apple.com/no/quicktime/

En annen måte å visualisere hvordan man oppfatter gestisk *chunking* er å bruke annoteringsprogrammet *Anvil*.⁶ *Anvil* er opprinnelig utarbeidet for å beskrive og kode non-verbal kommunikasjon, eller mer presist såkalte koverbale gester, dvs. det vi gjør med hendene mens vi snakker (Kipp 2004; McNeill 1992). Programmet gir mulighet til å spille av en videosekvens på den måten at denne kan stoppes hvor som helst på en tidslinje. Slik kan man markere start og stopp-punkt for en bevegelsesenhet, og man kan karakterisere eller merke enheten med tekst innenfor en ”tidsboks”.

Når det gjelder identifisering av musikalske *chunks* kan også QTP brukes til å klippe opp lydforløp. Men her vil jeg heller anbefale *Audacity*⁷ fordi man her får opp et bilde av lydets bølgeform. Visualiseringen er nyttig for å kunne ”holde fast” lyden tilstrekkelig til at man finner tilbake til det man hører som start og stopp-punkt for et segment. Museknappen brukes til å markere det området man ønsker å klippe ut. I figuren under er lydstrekket fra ca. 15 sek. til ca. 20 sek. markert slik at det kan klippes ut, og så eksporteres til en egen lydfil.



Figur 9. Audacity-skjermbilde

Å bryte opp lydstrømmen kan naturligvis også gjøres i et notebilde, men dette er sannsynligvis ikke så aktuelt for musikkterapeuter siden man sjelden har transkribert musikkutøvelsen til noter. Et annet poeng er at notebildet ikke nødvendigvis gir et nøyaktig bilde av tidsforløpet.

⁶ Programmet kan gratis lastes ned ved å henvende seg til Michael Kipp, som har laget det, se www.anvil-software.de/. Programmet kan brukes både på Mac og PC. Det har en liten terskel for å komme i gang, artikkelforfatteren er gjerne behjelpelig. Programmet er gratis.

⁷ Programmet kan enkelt lastes ned gratis, se <http://audacity.sourceforge.net/>

Beskrivelse av endringskarakteristikker – med utgangspunkt i aktiveringskontur

Etter å ha brutt opp den kontinuerlige strømmen av musikk og/eller bevegelse i mindre "chunks" er neste skritt å karakterisere disse segmentene mer detaljert. Som beskrevet tidligere foreslår jeg at chunks er grunnleggende karakterisert ved måten de endrer seg på innenfor tidsvinduet. Et musikk-segment kan f.eks. karakteriseres av endringer i tonehøyde, klang, og/eller artikulering. Tilsvarende kan en kroppsbevegelse karakteriseres gjennom endringer i tempo, utstrekning i rommet, involvering av kroppsdel, og/eller artikulering.

Her skal jeg ta utgangspunkt i hvordan endringer kan beskrives med hensyn til aktivering, dvs. endring i form av aktiveringskonturer. På grunnlag av de første "observasjonene" (lytte flere ganger til musikksporet/se mange ganger på videosporet), kan musikk og bevegelse for eksempel karakteriseres ved å si at "aktivering endrer seg fra mye til lite" – eller "fra mye til mindre, til mye igjen for så å avta til lite," osv. Observasjonene kan suppleres/kombineres med software-baserte representasjoner.

For bevegelsen vil en såkalt "Quantity-of-Motion" (QoM) analyse være aktuell. Kort beskrevet baserer analysen seg på registrering av forandring i pixlene fra bilderamme til bilderamme i en digital videostrøm. Dersom det er mye bevegelse (store deler av kroppen involvert og hurtige bevegelser), vil relativt stor grad av forandring i et stort antall pixler gi et høyt utslag og registreres som en høy grad av bevegelsesmengde. Når analysen gjøres over et tidsrom på noen sekunder, vil man se endringer i bevegelsesmengde som en funksjon av tid. I figur 10 er endringene plottet slik at man får en graf. Analysen er her gjort på den danse-varianten der bevegelsen starter med to påfølgende bølgelignende armbevegelser (se figur 8).⁸



Figur 10. QoM-analyse av bevegelsessekvens med økt aktivering mot slutten.

Grafen antyder at bevegelsene er rolige til å begynne med; at det er små endringer i aktivitetsnivå og at endringene er gradvis introdusert (avrundete kurver). Mot slutten skjer det en større endring i bevegelsesmengde, og denne

⁸ Analysen er gjort med "Musical Gestures Toolbox", utviklet av Alexander Refsum Jensenius, se <http://musicalgestures.uio.no>

inntreffer relativt abrupt (bratt kurve), noe som tyder på at bevegelsen vil oppfattes med en økning i aktivering her.

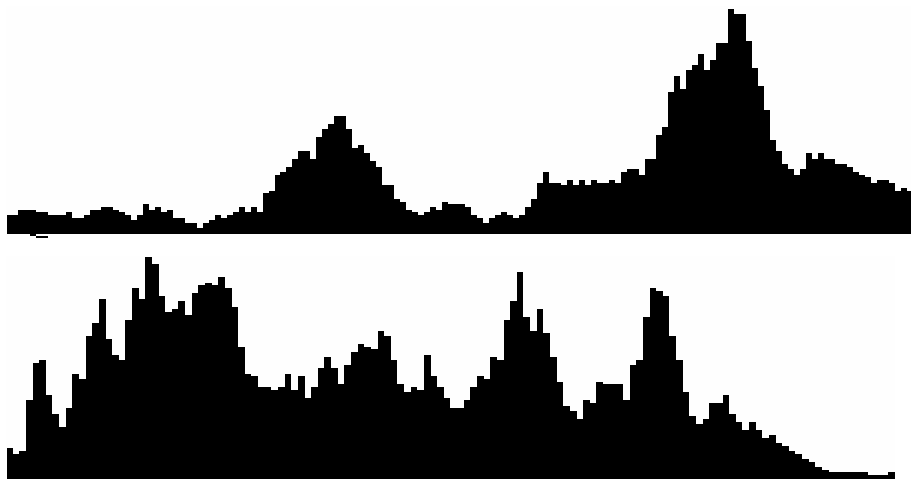
Slike analyser kan ikke leses for ”bokstavelig”. En beregning og representasjon av et video-signal er prinsipielt ikke det samme som en gjengivelse av perseptuell opplevelse, i dette tilfellet en aktiveringskontur. Mer konkret er det aspekter ved bevegelse som er usynlig for et video-kamera/datamaskin, men som vi er svært sensitive til. Jeg sikter her for eksempel til om bevegelsen er kontrollert eller ”sluppet”, eller om den er utført med tyngde eller letthet.⁹ Dette er aspekter knyttet til bevegelsens intensjonalitet, som sannsynligvis vil påvirke vår oppfattelse av aktiveringsendringer, men som den maskinelle analysen ikke vil fange opp fordi den måler kun rent fysiske aspekter ved bevegelsen.

Det er også en mer teknisk grunn til at analysen bør tas med en klype salt. Den datamaskin-baserte beregningen er ”dum” i den forstand at den registrerer alle pixel forandringer, ikke bare de som er relatert til den bevegelsen vi vil analysere. Hvis det er uroligheter i bakgrunnen (som kan skyldes video-opptaket, lysforhold), vil disse også registreres. Dersom bevegelsen av en arm bryter en visuell linje (en bordkant, overgangen fra gulv til vegg o.l.) i bakgrunnen vil dette påvirke graden av pixel-forandring, en endring som er skapt av andre faktorer enn bevegelsen selv. For at analysen skal bli så riktig som mulig, bør opptaksforholdene (godt lys, ikke fluoriserende) være gode og ideelt gjøres mot en ensfarget bakgrunn.¹⁰

På tross av disse innvendingene mener jeg det er verdt å bruke slike verktøy. De gir et bilde som er omtrent like godt eller like dårlig som andre framstillinger av et aspekt ved fenomenet. Ikke minst gir analysen en framstilling av endringer i forhold til en tidslinje og gir en ramme for videre diskusjon og beskrivelse av hvordan disse endringene inntreffer. Analyseverktøyet gir også en mulighet til å drøfte forskjeller mellom varianter. I figur 11 er to andre slike analyser. Den øverste er av en bevegelses-variant som er en tolkning av det samme musikkutsnittet som i det første eksemplet på en QoM-analyse (figur 10). Vi ser at endringen i bevegelsesmengde viser likheter (aktiveringsnivået er lavt til å begynne med og øker mot slutten). I den nederste er analysen av en bevegelse som er utført på grunnlag av et musikkutsnitt der aktivering er relativt høy fra starten for så å avta mot slutten. Denne endringen fra høy til lavere aktivitet reflekteres i analysen. De to kurvene i figuren er QoM-analyser av bevegelsesforløp som er basert på to ulike musikkutsnitt fra Gyögy Ligeti's *Ten Pieces for Wind Quintet*, det første fra en rolig sats, *Lento*, og det andre fra en hurtig sats, *Prestissimo*.

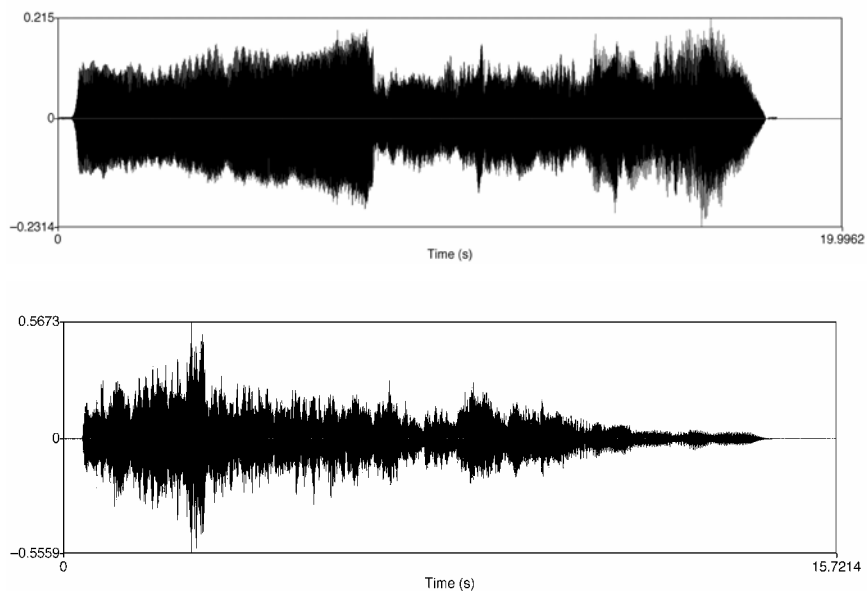
⁹ Disse aspektene ved en bevegelse er forsøkt fanget inn i *Laban Motion Analysis* (Laban 1971).

¹⁰ Disse problemene kan bøtes på ved å konvertere video-filen til et mp4-format som til en viss grad utligner uro i bildet.



Figur 11. QoM-analyse av to bevegelsesvarianter.

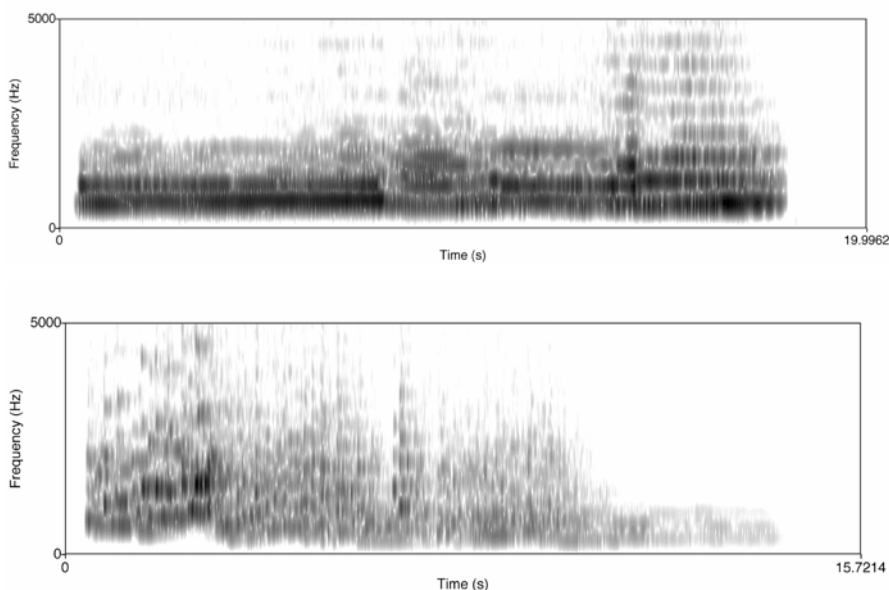
Tilsvarende kan man gi foreløpige antydninger av aktiveringskontur ved å analysere lydsignalet direkte. En *bølgeform* viser amplitude variasjoner som en funksjon av tid, og gir slik et grovt bilde av hvordan *loudness* endrer seg innenfor tidsvinduet (McAdams, Depalle & Clarke 2004). Figur 12 viser *bølgeform* av de to musikkutsnittene, *Lento* og *Prestissimo*.¹¹



Figur 12. Bølgeform-representasjon av to musikkutsnitt.

¹¹ Analysen er gjort med gratisprogrammet *Praat*, se www.fon.hum.uva.nl/praat/

Et *spektrogram* viser hvilke frekvenskomponenter som til en hver tid er tilstede i en lyd, og hvordan hver frekvens-komponent varierer i styrke over tid, dvs. ved sterkere gråtone (mot svart) er frekvenskomponenten sterkere. Figur 13 viser spektrogram av de samme musikk-utsnittene som over, henholdsvis *Lento* og *Prestissimo*. Spektrogrammet supplerer bølgeformen ved å si noe om den vertikale distribusjonen av frekvenskomponenter. Sammen med amplituden av et lydsignal spiller det en rolle for hvilken styrke vi oppfatter signalet med hvilke frekvenser som er framtrepende, fordi øret er mer sensitivt i noen frekvensområder enn andre (Mathews 2001). I analysen av *Lento*-utsnittet er det sterkere frekvenser i et slikt sensitivt område mot slutten av forløpet. Dette indikerer at lyden vil oppfattes som mer påtrengende her, noe som sannsynligvis vil oppfattes med en høyere grad av aktivering.



Figur 13. Spektrogram av to musikkutsnitt.

En annen fordel med å bruke slike maskinelle analyser av et musikkutsnitt er at de gir en nøyaktig tidslinje som man kan ta utgangspunkt i når man skal gå videre med å kommentere på mer kvalitativt grunnlag.

Aktiveringstrekk i musikk og bevegelse

Siden analysen altså ikke sier noe om hvordan endring skjer, skal jeg i det følgende foreslå en måte å beskrive endringen mer detaljert. Igjen med utgangspunkt i aktiveringskontur som overordnet beskrivende term, antar jeg at det er mulig å fokusere på, eller skille ut, noen trekk i henholdsvis musikk og gestikk som til sammen bidrar til oppfattelsen av aktivering.

Dette bygger på en grunnleggende forståelse av at musikk, og bevegelse, er ”satt sammen” av mange enkeltelementer som til sammen gir opphav til den helheten vi oppfatter med sansene våre. Denne helheten forstås som en *global emergent egenskap*, dvs. en egenskap som trer fram for sansene våre som en overordnet egenskap. *Aktiveringskontur* er eksempel som en slik global egenskap som trer fram fordi flere elementer virker sammen og påvirker hverandre (andre eksempler på *globale emergente egenskaper* fra musikk er klang og rytme). Det er viktig å merke seg at den globale egenskapen ikke kan forutsies som i et regnestykke; elementene summeres ikke enkelt sammen, men griper ofte inn i hverandre og påvirker hverandre gjensidig. I det følgende vil jeg omtale enkeltelementene som medvirker til oppfattelse av aktivering i henholdsvis musikk og bevegelse som *aktiverings-trekk*.

For musikkens del vil jeg anta at følgende *aktiverings-trekk* er medvirkende:

- *Tetthet av hendelser*, dvs. høyere tetthet (flere hendelser innenfor et tidsvindu) fører til oppfattelse av en høyere grad av aktivering. Med hendelser menes tone- og akkordansatser.
- *Tone-høyde endringer*, dvs. relativt store endringer i tonehøyde gir en opplevelse av høyere grad av aktivering.
- *Loudness*, dvs. musikk spilt med en styrkegrad som i fortissimo fører til oppfattelse av en høyere grad av aktivering.
- *Artikulering*, dvs. musikk som er distinkt/markert/skarpt artikulert (som i staccato/marcato) antas å oppfattes med en høyere grad av aktivering enn musikk som er mykere artikulert (som i legato).
- *Timbre/klangfarge*, dvs. en påtrengende/intens klang antas å bli oppfattet med en høyere grad av aktivering enn en klangfarge som er rundere/mykere.

Tilsvarende foreslår jeg at følgende *aktiverings-trekk* medvirker til oppfattelse av aktivering i bevegelser:

- *Tetthet av hendelser i tid*, dvs. hyppigheten i tid av retningsendringer, aksentueringer i en bevegelse, og start-punkter for nye *chunks*. Høyere tetthet fører til høyere grad av aktivering.
- *Utstrekning av bevegelser*, dvs. bevegelser med stor utstrekning i rommet vil lede til en høyere grad av aktivering.
- *Involvering av kroppsdel*, dvs. bevegelser som involverer store deler av kroppen vil bli oppfattet med økt grad av aktivering.
- *Tempo*, dvs. hurtige bevegelser vil resultere i en høyere grad av aktivering.
- *Kraft*, dvs. kraftfulle bevegelser vil føre til en høyere grad av aktivering.
- *Artikulering*, dvs. ’skarpt’ aksentuerte bevegelser som ved abrupte endringer i retning, vil oppfattes med en høyere grad av aktivering.
- *Betoninger*, dvs. bevegelser der det er en høy frekvens av betonte øyeblikk, vil oppfattes med en høyere grad av aktivering.

Aktiverings-trekkene for henholdsvis musikk og bevegelse som jeg har skissert her, må forstås som antagelser om “hva” i musikken som gjør at vi hører endringer i aktivering, og antagelser om “hva” i bevegelsen som gjør at vi ser endringer i aktivering. De er ment som observasjonskategorier som man kan bruke til å beskrive endring på et mer detaljert nivå.

Framgangsmåten for analysen vil være at man lytter gjennom musikken flere ganger og fokuserer på ett og ett trekk og spør seg:

- Hvordan endrer hendelsestetthet seg i dette utsnittet?
- Hvordan endrer tonehøyde seg i dette utsnittet?
- Hvordan endrer loudness seg i dette utsnittet?
- Hvordan endrer artikulering seg i dette utsnittet?
- Hvordan endrer klangen seg i dette utsnittet?
- Hvordan endrer teksturen seg i dette utsnittet?

På bakgrunn av gjentatte lyttinger kan en beskrivelse av endring i de ulike aktiverings-trekkene eksempelvis se ut som figur 14, der endringene er stilt opp sammen med spektrogrammet:

<i>Tetthet hendelser</i>	Lav	Mer	Mindre
<i>Tonehøyde variasjon</i>	Ingen	Mer	Trinnvis
<i>Loudness</i>	Lite (pp)		Svakt økende
<i>Artikulering</i>	Alle toneinnsatser er 'myke' (ingen endring)		
<i>Klang</i>	Lys	Noe lysere/mer intens	
<i>Tekstur</i>	Tett	Åpner opp/sprer seg mer	
<i>Spektrogram med oppfattet aktiveringskontur inntegnet</i>			

Figur 14. Endringer i aktiveringstrekk, kvalitativt vurdert.

For å notere disse observasjonene er *Anvil* et nyttig verktøy. Her skal jeg vise hvordan programmet kan brukes for å notere endring i *aktiverings-trekk* i et

bevegelsesforløp. Annoterings-plansjen settes opp slik at hvert aktiveringstrekk er representert.¹² Dette er vist i figur 15.

activation	density	
	extension	
	involvement	
	speed	
	force	
	articulation	
	emphasis	

Figur 15. Aktiverings-trekk i et bevegelsesforløp notert inn i *Anvil*.

Her er plansjen satt opp slik at:

- Hver tidsboks "kodes" med 'mye/moderat/lite' (alternativt 'rask/moderat/langsom').
- Og slik at disse grads-forskjellene visualiseres med farger.

Resultatet er at gradsforskjeller vises langs en tidslinje, der start og slutt punkt for endring/nivå angis med en tidsboks. På denne måten visualiseres endring. I eksemplet under har jeg notert hvordan jeg, basert på gjentatte observasjoner, oppfatter endringer i aktiverings-trekk i et bevegelses-forløp (den samme bevegelsesvarianten som er vist til tidligere, dvs. den som starter med to bølgelignende armbevegelser, se figur 8). Bevegelsene er generelt rolige og 'uttrukne', med en svak økning i aktivering mot slutten. Dette er illustrert med endring i gråtoner. Plansjen er sammenstilt med QoM-analysen for det samme forløpet. Til sammen gir dette et godt bilde av hvordan man kan oppfatte aktivering i sekvensen, både mht. overordnet kontur ('lite – litt mer – til avtagende') og hvor på tidslinjen denne endringen skjer i de ulike aktiverings-trekkene.

Aktivering	density	low	moderate	l..
	extension	low	moderate	low
	involvement	low	moderate	low
	speed	low	moderate	low
	force	weak		
	articulation	smooth		
	emphasis			

¹² Utseende og egenskaper til annoterings-plansjen gjøres i xml-kode.

QoM-
analyse



Figur 16. (jf. med slutten av forrige side) Kvalitativ vurdering av endringer i aktivering sammenstilt med QoM-analyse.

Avrunding

De erfaringene jeg bygger på er gjort med utgangspunkt i video-opptak av trente dansere som ble bedt om å bevege seg til musikkutsnitt med kort varighet; videre at valget av musikkstil er begrenset til ikke-periodisk musikk uten et klart definert metrum (2/4, 3/4 eller lignende) og en "beat". I musikkterapeutisk praksis vil jeg anta at en "beat" eller et rytmisk mønster ofte vil være tilstede og ha en innrammende funksjon i samspillet. Jeg tror allikevel at fokuset jeg har foreslått vil kunne tilpasses og anvendes.

Når de gjelder de software-baserte verktøyene jeg har foreslått er det forskjeller mht. tilgjengelighet og brukervennlighet. QTP, Audacity og Praat er alle enkle å bruke, enkle å laste ned og gratis (bortsett fra QTP der lisensen koster et mindre beløp). Anvil er også gratis men krever en god del mer for å komme i gang med (tilpasning av annoteringsplansje ved hjelp av Xml-kode er bøygen for de som ikke har erfaring fra før). Musical Gestures Toolbox som QoM-analysene er laget med, er fritt nedlastbart, relativt lett å bruke, men krever at man har programmet Max/MSP/Jitter installert (og det er dyrt).

Litteratur

- Bregman, A.S. (1990). *Auditory Scene Analysis. The Perceptual Organization of Sound*. Cambridge, Mass.: The MIT Press.
- Clarke, E. F. (2005). *Ways of Listening. An Ecological Approach to the Perception of Musical Meaning*. Oxford: Oxford University Press.
- Deutsch, D. (1999). *The Psychology of Music* (2nd ed.). New York: Academic Press.
- Feldman, J., Epstein, D. & Richards, W. (1992). Force Dynamics of Tempo Change in Music. *Music perception*, 10 (2). 185-203
- Kendon, A. (2004). *Gesture. Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kipp, M. (2004). *Gesture Generation by Imitation - from Human Behavior to Computer Character Animation*. Dissertation.com.
- Laban, R. (1971). *The Mastery of Movement*. Plymouth: MacDonald and Evans.
- Lewkowicz, D. J., & Kraebel, K. S. (2004). The Value of Multisensory Redundancy in the Development of Intersensory Perception. I: Calvert, G.A.,

- Spence, C. & Stein, B. E. (red.) *The Handbook of Multisensory Processes*. Cambridge, Mass.: MIT Press.
- Lewkowicz, D. J. & Turkewitz, G. (1980). Cross-Modal Equivalence in Early Infancy: Audio-Visual Intensity Matching. *Developmental psychology*, 16, 597-607.
- Mathews, M. (2001). What Is Loudness. I: Cook, P.R. (red.), *Music, Cognition and Computerized Sound*. Cambridge, Mass.: MIT Press.
- McAdams, S., Depalle, P. & Clarke, E. (2004). Analyzing Musical Sound. I: Clarke, E. & Cook, N. (red.), *Empirical Musicology*. Aims, Methods, Prospects. Oxford: Oxford University Press.
- McNeill, D. (1992). *Hand and Mind. What Gestures Reveal About Thought*. Chicago: Chicago University Press.
- Meltzoff, A. N. & Borton, W. (1979). Intermodal Matching by Human Neonates. *Nature*, 282, 403-404.
- Rosenbaum, D. A. (1990). *Human Motor Control*. San Diego, Cal.: Academic Press.
- Runeson, S. & Frykholm, G. (1983). Kinematic Specification of Dynamics as an Informational Basis for Person-and Action Perception: Expectation, Gender Recognition and Deceptive Intention. *Journal of experimental psychology, General*, 112, (4), 585-615.
- Snyder, B. (2000). *Music and Memory: an introduction*. Cambridge, Mass.: MIT Press.
- Stern, D. (2000). *The Interpersonal World of the Infant*. New York: Basic books.
- Todd. (1995). The Kinematics of Musical Expression. *Journal of the acoustical society of America*, 97(3). 1940-1949
- Trondalen, G. & Skårderud, F. (2007). Playing with affects. *Nordic Journal of Music Therapy*, 16 (2), 100-112.
- Winter, D. A. (2005). *Biomechanics and Motor Control of Human Movement*. Hoboken, N.J.: Wiley. 3rd ed.