

# Here, there and everywhere

*Konsumenters opplevelse av tidlige refleksjoner i reprodusert populærmusikk*

**Lars Svanberg Jakobsen**



Hovedfagsoppgave i musikk

Institutt for musikkvitenskap

**UNIVERSITETET I OSLO**

November 2006

## Forord

Det er flere som må takkes for sitt bidrag til denne oppgaven.

Først en takk til veileder Tellef Kvifte og 90-tallsveiledere Rolf Inge Godøy og Stan Hawkins. Takk til NOTAM (Norsk nettverk for Teknologi, Akustikk og Musikk) for lån av utstyr og tilgang til studiofasiliteter, og Jørn Rune Kviserud som velvillig lånte ut sitt ekkofrie rom.

Takk til Lise Bugge-Amundsen som bidro med vakker sang i ekkofrie omgivelser. Takk også til Kristin Benestad, Inger Margrete Eggen, Torgeir Knag Fylkesnes, Audun Herning og Lars Kolltveit for at de ville delta i lytteforsøket. Med unntak av regissør og arkeolog Eggen, er de andre gode kollegaer ved min arbeidsplass på partikontoret til Sosialistisk Venstreparti.

En stor takk til Jonas Wille og Ingrid Nakken som begge har bidratt i både forforsøk og lytteforsøk, og som gav meg en stor opptur midt i skivinga ved å la meg knuse dem i FIFA 😊

Aller mest vil jeg takke min kjære samboer Elin som har bidratt stort i både forsøk og tilværelsen ellers. Sist og minst, en stor klem til Nora og han/hun som vokser inni magen til Mamma.

Oslo, 27. oktober 2006

Lars Svanberg Jakobsen



# Innhold

1	Innledning.....	1
1.1	Utdyping av problemstillingen.....	3
1.2	Konkretisering.....	7
1.2.1	Kommunikasjon med deltakeren.....	9
1.2.2	Hypoteser.....	9
1.3	Teoretisk fundament.....	10
1.4	Metode.....	11
1.5	Forutsetninger hos leseren.....	14
1.6	Avgrensing.....	15
1.7	Begrepsavklaring.....	16
2	Tidlige refleksjoner.....	23
2.1	Hvordan opplever vi romklang.....	23
2.2	Akustikken.....	25
2.2.1	Hva er tidlige refleksjoner.....	25
2.2.2	En enkelt refleksjon.....	26
2.2.3	Kamfiltre og stående bølger.....	27
2.2.4	Flere refleksjoner.....	29
2.2.5	Antall tidlige refleksjoner.....	30
2.3	Psykoakustikken.....	30
2.3.1	Lydkilder.....	31
2.3.2	Romopplevelser.....	34
2.3.3	Grenseverdier.....	36
3	Musikkproduksjon.....	38
3.1	Innspillingen.....	38
3.1.1	Lydkildens posisjon og karakteristik.....	41
3.1.2	Mikrofonens posisjon og karakteristik.....	42
3.1.3	Kontrollrom.....	45
3.2	Effektprosessering/miksing.....	46
3.3	Kunstig romklang.....	48
3.3.1	Naturlig og unaturlig romklang.....	49
3.3.2	Algoritmebaserte klanger.....	51
3.3.3	Impulsrespons.....	55
3.4	Bruken i praksis.....	56
4	Konsumentens lyttesituasjon.....	59
4.1	Effekten av ulike lyttesituasjoner.....	60
4.2	Nærhet til vegger.....	61
4.3	Dimensjonene.....	62
4.3.1	Høyde.....	62
4.3.2	Bredde.....	63
4.3.3	Dybde.....	64
4.4	Typiske trekk ved konsumenters lytting.....	67
4.4.1	Høytalerplassering.....	68
4.4.2	Begrenset båndbredde.....	69
4.4.3	Avspillingsvolum.....	69
5	Beskrivelse av lytteforsøk.....	71
5.1	Innspilling i Rakkestad.....	71
5.1.1	Det ekkofrie rommet.....	72

5.1.2	Innspillingsutstyr.....	72
5.1.3	Musikere og instrumenter.....	73
5.2	Miksing.....	73
5.2.1	Lydeksemlene.....	74
5.2.2	Romklngen .....	76
5.2.3	Impulsresponser .....	78
5.2.4	Lydstyrken på de tidlige refleksjonene .....	79
5.2.5	Det som ikke ble gjort .....	79
5.2.6	Nivåjustering .....	79
5.2.7	Feil på nivåer i grensesnittet.....	81
5.3	Grensesnittet.....	81
5.3.1	Flash-applikasjon .....	81
5.3.2	Graderingsskala.....	82
5.4	Oppsett i lytterom.....	82
5.4.1	Høytalerplassering.....	83
5.4.2	Lytterplassering .....	83
5.4.3	Impulsresponser av posisjonene.....	84
5.4.4	Avspillingsnivå.....	85
5.5	Gjennomføring av forsøket .....	85
5.5.1	Oppgavene.....	85
5.5.2	Forsøkspersoner .....	86
5.5.3	Instruksjoner til deltakeren.....	86
5.5.4	Intervju .....	87
6	Analyse.....	88
6.1	Begreper .....	89
6.2	Hypotesene .....	90
6.2.1	Hypotese 1 .....	90
6.2.2	Hypotese 2.....	91
6.2.3	Hypotese 3.....	92
6.2.4	Hypotese 4.....	92
6.2.5	Hypotese 5.....	93
6.3	Feilprosent.....	93
6.4	Tidsbruk .....	95
6.5	Bruk av skalaen .....	96
6.5.1	Lik gradering .....	98
6.6	Gitaren.....	99
6.7	Noen observasjoner .....	100
6.8	Mulige feilkilder.....	101
6.9	Avsluttende betraktninger .....	102
7	Konklusjon .....	105
	Litteraturliste .....	111
	Appendiks A.....	117
	Appendiks B.....	121

# 1 Innledning

Arbeidstittelen til denne oppgaven ble formulert for hele ti år siden og har hele tiden vært ”Bruk av kunstig romklang i populærmusikkproduksjon”. Bakgrunnen for valget av arbeidstittel lå i et ønske om å belyse hvordan lytterens musikkopplevelse påvirkes av de valgene produsenter og lydteknikere gjør med hensyn til bruk av kunstig romklang. Arbeidet har utviklet seg mye på den tiden, men det er fremdeles med dette ønsket i bakhodet jeg har endt opp med å formulere problemstillingen i oppgaven.

Problemstillingen i oppgaven er som følger: *”Hvordan oppfatter konsumenter spatiale trekk i lydbildet til populærmusikkproduksjoner reprodusert i stereo? I tilnærmingen til problemstillingen rettes hovedfokus på tidlige refleksjoner og ikke-optimale lyttesituasjoner.”*

Problemet med hovedoppgaver er at de kan virke så uendelig lange å fullføre, men har allikevel så uendelig lite plass når man skal velge fokus for oppgaven. Følelsen forsterkes ytterligere når man ønsker å belyse et emne som er lite behandlet fra før. Som man forstår er den endelige problemstillingen kun en liten del av det store feltet jeg i utgangspunktet ønsket å belyse. Det er naturlig å anta at ønsket om en helhetlig framstilling har farget mitt arbeid med oppgaven. Oppgaven bør således også leses med dette i bakhodet.

Musikkopplevelse er fremdeles et viktig stikkord for oppgaven. Musikkopplevelse foregår både på det bevisste og det ubevisste plan, og det er distinksjonen mellom disse planene sett i forhold til konsumenter jeg håper å belyse med problemstillingen. Jeg kommer tilbake til en grundigere definisjon av konsumenten, men den store majoriteten av psykoakustisk, akustisk og ikke minst musikkvitenskapelig arbeid som omhandler opplevelse av musikkens spatiale trekk, tar utgangspunkt i den skolerte lytter. De fleste arbeider som berører temaet omhandler klassisk musikk i reprodusert eller levende form. Der er muligens

dannelsen høyere, men den store, store majoriteten av lyttere til populærmusikk er ikke skolerte i emnet. Så mens Timothy Warner opplever Trevor Horns kastanjetter rett i trynet (Warner, 2003), lurer jeg på om Trine på tretten opplevde annet enn vann i ansiktet da hun sto på kjøkkenet og vasket opp mens hun nynet til Grace Jones. Jeg mener på ingen måte å diskvalifisere verken Trine eller Warner. Sistenevntes fokus er også et helt annet, men det er et godt eksempel på spatiale musikkopplevelser jeg tror de færreste lyttere opplever bevisst.

Den andre setningen i problemstillingen retter fokus inn mot to områder. Fokuset på tidlige refleksjoner er et tiltak for å snevre inn nedslagsfeltet til problemstillingen, mens fokuset på ikke-optimale lyttesituasjoner muligens utvider det. Jeg skal si litt om bakgrunnen for valget av begge.

Romklang er et annet navn på lydbølger som på sin vei fra lydkilde til lytter har vært innom en eller flere overflater. Med andre ord, romklang er refleksjoner. I akustikken deler man romklangen inn i henholdsvis tidlige og sene refleksjoner, hvor de tidlige refleksjoner er de første refleksjonene som når lytteren. Kapittel 2 beskriver dette ytterligere. I populærmusikkproduksjoner benytter man gjerne kunstig romklang til å simulere akustiske omgivelser. Kunstig romklang blir definert i begrepsavklaringen senere i kapittelet, men som navnet innebærer skapes den på kunstig vis, i kontrast til den naturlige romklangen. Kunstig romklang vært fokus for arbeidet med hovedoppgaven fra starten av. Undertegnedes ønske om å beholde et fokus på romklang er derfor naturlig nok til stede. Spatiale aspekter i populærmusikkproduksjoner kan deles inn i mange dimensjoner, alt for mange for en liten hovedoppgave. Hovedfokuset på tidlige refleksjoner blir dermed også et tiltak for å snevre inn oppgavens fokus ytterligere. Et hovedfokus på romklangens tidlige refleksjoner betyr selvsagt ikke at jeg kan overse alle andre spatiale faktorer.

Det er flere årsaker til ønsket om å belyse ikke-optimale lyttesituasjoner. Med bakgrunn i alle kriterier som er avgjørende for en optimal lyttesituasjon for reproduksjon av musikk i stereo, er det naturlig å anta at majoriteten av konsumenter sjelden befinner seg i en slik situasjon. Ulikheter i lyttesituasjon er et tema som sjelden berøres av musikkvitenskapen. Blant annet er lytterens posisjon i forhold til høyttalerne helt kritisk for å oppnå korrekt gjengivelse av stereo musikk. Selv innen akustikken og musikkteknologien synes posisjoner borte fra en optimal posisjon å være et neglisjert område. Fokuset har riktig nok endret seg litt i takt med surround-anleggenes økende popularitet, men det savnes fremdeles utdyping av reproduksjon i stereo. Stereomediet er fremdeles tilnærmet enerådende innen musikkformidling, og vil sannsynligvis være det i noe tid framover. Mitt bidrag blir lite, men

fokuset på ikke-optimale lyttesituasjoner tvinger seg fram når man skal belyse konsumenters lytting til musikk.

Det videre forløpet i oppgaven er lagt opp på følgende måte. I fortsettelsen av dette kapittelet vil jeg starte med å utdype problemstillingen ved å definere begrepene som er benyttet i denne. Beskrivelsen av problemstillingen etterfølges av en konkretisering. Konkretiseringen beskriver hvordan jeg har tenkt å belyse problemstillingen gjennom et praktisk lytteforsøk. En konsekvens av konkretiseringen i form av et praktisk lytteforsøk, innebærer en ytterligere innsnevring av fokus for oppgaven. I fortsettelsen av kapittelet beveger jeg meg inn på det teoretiske fundamentet for oppgaven, og fortsetter med valget av metode. Jeg avslutter med hva som forutsettes av leseren, en avgrensning av omfanget og helt til slutt en generell begrepsavklaring.

Innholdet i oppgaven er delt inn i 7 kapitler. Innledningen utgjør første kapittel. Kapittel 2 er viet de tidlige refleksjonene og hvordan disse spiller en rolle i lyden vi hører. Tilnærmingen baserer seg på de grunnleggende premissene som ligger i akustikken og psykoakustikken. Kunnskapen fra kapittel 2 leder oss videre til kapittel 3, som ser på lydstudioet der musikkproduksjonen foregår. Kapittelet starter med et blikk på hvordan de naturlige skapte tidlige refleksjonene preger både det innspilte materialet og lytteforholdene til lydtekniker/producent. Resten av kapittel 3 er viet den kunstig skapte romklangen og hvordan denne skapes og brukes i konstruksjonen av populærmusikkproduksjonene vi hører. Kapittel 4 tar tak i noen av de ikke-optimale lyttesituasjonene som preger konsumentens lytting til reproduisert musikk. Kapittelet forsøker å vise hvordan vår opplevelse av musikken påvirkes av rommet der musikken reproduseres, og ikke minst lytterens posisjon i dette rommet. Kunnskapen som presenteres i kapittel 2 til 4 danner utgangspunktet for det praktiske lytteforsøket som er gjennomført i forbindelse med oppgaven. Kapittel 5 beskriver forberedelsene og gjennomføringen av dette forsøket, mens kapittel 6 er viet til analysen av deltakernes besvarelser i forsøket. Oppgaven avsluttes med en oppsummering og et blikk fremover i kapittel 7.

## **1.1 Utdyping av problemstillingen**

I de følgende avsnittene vil jeg utdype begrepene som er benyttet i problemstillingen, og håper på den måte å få fram tanken bak problemstillingen. Beskrivelsen er strukturert som en begrepsavklaring, men må ikke forveksles med den mer generelle begrepsavklaringen som kommer i kapittel 1.7. Til hjelp for leseren, kommer her en repetisjon av problemstillingen:



*”Hvordan oppfatter konsumenter spatiale trekk i lydbildet til populærmusikkproduksjoner reproduisert i stereo? I tilnærmingen til problemstillingen rettes hovedfokus på tidlige refleksjoner og ikke-optimale lyttesituasjoner.”*

### *Oppfatte*

Å oppfatte er et tvetydig begrep. Noe av tvetydigheten trer godt fram når man beveger seg inn på områder som hørsel, psykoakustikk og musikkopplevelse. Akkurat det gjør ordet egnet til min problemstilling. Ordboka sier at ordets opprinnelige betydning er å ”gripe (med tanken)” (”Bokmålsordboka”, 2006). Ordboka lister deretter opp tre ulike betydninger av ordet: å høre, å forstå og å tolke. Alle tre er gyldige for problemstillingen.

Å gripe noe med tanken bør gjerne innebære en form for bevissthet, eller erkjennelse om man vil, men det er allikevel ikke uvanlig å kople ordene ”ubevisst” og ”oppfatte” sammen. Man finner mange eksempler på dette i blant annet film- og reklameverden, hvor man ved produktplasseringer og symbolbruk arbeider for å påvirke oss forbrukere. Man foretrekker kanskje å benytte uttrykket underbevisst framfor ubevisst i slike sammenhenger, men ordene benyttes gjerne med lik betydning (”Caplex”, 2006).

Kombinasjonen ubevisst og oppfatte er like aktuell når det gjelder hørselen og våre opplevelser av lyden som treffer trommehinnene våre. Hørselen analyserer og identifiserer veldig mye av lyden den mottar, men det aller meste forblir i underbevisstheten. Det blir problematisk å hevde at man ikke oppfatter en passerende bil, selv om lyden av bilen aldri når opp til bevisstheten vår. Et mer passende eksempel i denne sammenheng er vår fantastiske evne til å identifisere romlige omgivelser ut fra lyden vi hører. Hvor ofte når denne identifiseringen bevisstheten til en som ikke studerer emnet? Svaret er antakeligvis relativt sjelden. Det er her kjernen i min problemstilling ligger. Spørsmålet er i hvilken grad spatiale trekk som hjernen formodentlig identifiserer, altså oppfatter, når opp til lytterens bevissthet.

I forlengelsen av dette, må det nevnes at ordet bevissthet bringer inn mange nye spørsmål rundt problemet. Når kan man for eksempel si at en lytter er bevisst på et spatialt trekk i lydbildet? Er det i det man kan identifisere det spatiale trekket og sette ord på det, eller holder det at man legger merke til at noe er annerledes sammenlignet med et annet lydbilde? Problemstillingen tar ikke mål av seg å belyse disse, eller andre spørsmål som faller inn under den kognitive sfæren. Den observante leser vil etter hvert merke seg at jeg i den praktiske delen i oppgaven tar sikte på å avdekke i hvilken grad konsumenter evner å identifisere et spesifikt spatialt trekk, og således har tatt stilling til spørsmålene ovenfor.

### *Konsument*

Konsumentene er ingen entydig gruppe. I utgangspunktet kan antakeligvis hele befolkningen klassifiseres som konsumenter av musikk. Min definisjon begrenser seg til noen få kriterier.

1. Konsumenten lytter til innspilt musikk via stereoanlegg, radio eller tv.
2. Konsumenten er ikke involvert i produksjon av musikk.
3. Konsumenten er ikke skolert/trent i lydarbeid, lydteknikk eller annen aktivitet som innebærer kritisk lytting.
4. Konsumenten har binaural hørsel (binaural = to ører).

Jeg forsøker her å skape et bilde av den jevne mann i gata. En som ikke er opptatt av lyd i utgangspunktet, men som med fordel er opptatt av musikk. Det man skal merke seg er at jeg til en viss grad forsøker å utelukke såkalte audiofile. Det kan synes merkelig å utelukke en relativt stor gruppe mennesker fra en kategori de ubønhørlig hører hjemme i. Det karakteristiske ved audiofile er at de tilstreber å oppnå en optimal lyttesituasjon. Det er de konsumenter som ikke gjør det, jeg er opptatt av.

### *Spatial*

Definisjonen på spatial hentes rett fra ordboka, og innebærer "som gjelder avstands- og romforhold" ("Bokmålsordboka", 2006). Det kan argumenteres at "romlig" er et begrep som både er bedre kjent og mer egnet, men dette er valgt bort fra problemstillingen for å knyttes tettere opp til spatiale aspekter som omhandler romklang.

### *Lydbilde*

Lydbilde er i denne sammenheng all lyd som formidles fra høyttalere og treffer lytteren. Lyd formidlet i stereo kan danne et lydbilde som formidler spatiale trekk i alle de tre dimensjonene vi lever i. Gjengivelsen i de tre dimensjonene er allikevel sterkt begrenset ved at man hovedsakelig danner et lydbilde i fremre kvadrant (Rumsey, 2001, s.53). Med fremre kvadrant menes området foran lytteren, når denne er vendt mot høyttalerne.

### *Populærmusikkproduksjoner*

Min bruk av populærmusikkbegrepet er av pragmatisk art og retter seg mer mot produksjonsmetode og -estetikk, enn mot en typisk genreklassifisering. Det er allikevel beleilig å ty til genrebegrep i en så grov avgrensing. Produksjonsmessig kan man grovt sett dele inn innspillinger av musikk i to kategorier, klassisk- og populærmusikkinnspillinger. Populærmusikkproduksjoner karakteriseres gjerne ved utstrakt bruk av flersporsopptak og ikke-simultan innspilling av instrumentene. I tillegg til innspillingsmetode er bruken av

effektprosessering et karakterisk trekk ved mange populærmusikkproduksjoner. Tilføring av kunstig romklang faller inn under begrepet effektprosessering. Det er veldig glidende overganger her, og kriteriet om ikke-simultan innspilling utelukker blant annet en stor del jazzinnspillinger. Det er ikke meningen. Det er ikke behov for vanntette skott i min definisjon, da det rent musikalske i og for seg ikke er i fokus.

#### *Reprodusert lyd*

Med reprodusert lyd menes lyd som har tatt veien om et lagringsmedium før det formidles gjennom høyttalere. Med andre ord, innspilt musikk.

#### *Stereo*

Stereo-begrepet bør være kjent for de fleste, men for ordens skyld sakser jeg følgende to definisjoner fra ordboka: *stereo*: (kortord av *stereofoni*) – *stereofoni*: elektroakustisk lydbølgeoverføring som gir romfølelse ved at lyden ledes gjennom to atskilte kanaler, til forskjell fra *monofoni* ("*Bokmålsordboka*", 2006).

#### *Tidlige refleksjoner*

Tidlige refleksjoner er kort beskrevet allerede. Det henvises til kapittel 2 for ytterligere utdyping.

#### *Ikke-optimal lyttesituasjon*

Jeg har valgt å fokusere på lyttesituasjoner som ikke er optimale. Det ligger i ordet at en ikke-optimal lyttesituasjon er noe som avviker fra en optimal lyttesituasjon. Det vil bli gitt en beskrivelse av den optimale lyttesituasjonen i begrepsavklaringen senere i kapittelet, men den forutsetter blant annet et presist avstandsforhold mellom høyttalere og lytter. Med helt klare rammer for lytterens posisjon, og i tillegg en rekke andre begrensende faktorer, er det naturlig å anta at den jevne konsument sjelden befinner seg i en optimal lyttesituasjon. Siden konsumenter er fokus for denne oppgaven, blir det også naturlig å fokusere på ikke-optimale lyttesituasjoner. Mitt fokus på ikke-optimale lyttesituasjoner vil rette seg inn mot lytteposisjoner som er noen meter borte, og til side for den optimale lytteposisjonen. I tillegg vil det også være et fokus på romlige forhold. Aspekter som sannsynligvis preger de fleste konsumenters musikklytting.

For å unngå forvirring, ber jeg leseren merke seg sammenhengen mellom begrepene lyttesituasjon og lytteposisjon, i kombinasjon med begrepene optimal og ikke-optimal. En optimal lyttesituasjon forutsetter blant annet at lytteren befinner seg i en optimal lytteposisjon, men den optimale lytteposisjonen er det ikke det eneste som konstituerer den optimale

lyttesituasjonen. Man kan dermed ha situasjoner der lytteren befinner seg i en optimal lytteposisjon, hvor den totale lyttesituasjonen er ikke-optimal.

## 1.2 Konkretisering

Problemstillingens ordlyd tar sikte på å avdekke konsumenters opplevelse av spatiale trekk i populærmusikkproduksjoner. Det er en ambisiøs problemstilling på flere måter. Studiet av menneskers musikkopplevelse er generelt et meget komplekst felt, uansett innfallsvinkel og fokus. Når jeg i denne oppgaven tar sikte på å belyse aspekter ved musikkopplevelsen som er lite utforsket fra før, er det stor fare for at jeg trækker feil på veien. Skulle jeg være så heldig å trække noenlunde riktig, kan man håpe at jeg klarer å tilføre noe meningsfylt til kunnskapen rundt dette komplekse emnet.

Det er spesielt to aspekter ved problemstillingen som gjør den kompleks. For det første tar den sikte på å si noe om musikkopplevelsen til uskolerte lyttere. Man må anta at den store majoriteten av befolkningen er uskolert i emnet. Det innebærer at problemstillingen søker å avdekke trekk som gjelder for et stort mangfold av mennesker. Problemstillingen søker også å si noe om musikkopplevelse i ikke-optimale lyttesituasjoner. Vi mennesker omgir oss med musikk i et stort mangfold av steder og situasjoner, hvor så godt som alle i større eller mindre grad kan beskrives som ikke-optimale. Med et mangfold som er så stort er det selvsagt utenkelig å skulle oppfylle problemstillingens mål i sin helhet. Når emnet i tillegg er relativt lite utforsket fra før, blir den eneste fruktbare innfallsvinkelen å begrense seg til en liten bit av den store helheten. Hovedfokus på romklangens tidlige refleksjoner begrenser selvsagt omfanget noe, men det har vært nødvendig å snevre inn enda mer. Behovet for ytterligere innsnevring åpenbarer seg i det man velger å belyse problemet gjennom praktiske forsøk.

Ønsker man å belyse konsumenters opplevelse av spatiale trekk i musikken, må man nødvendigvis forsøke å fange inn konsumentens opplevelse på et eller annet vis. Jeg har valgt å gjøre dette gjennom et praktisk lytteforsøk. Et grunnleggende trekk ved praktiske lytteforsøk, i alle fall av den typen som gjennomføres her, er å tilstrebe en lavest mulig grad av usikre faktorer. En effektiv måte å begrense usikre faktorer på, er å snevre inn fokus.

Tidlige refleksjoner kan være en medvirkende faktor i opplevelsen av en lang rekke spatiale trekk. Jeg valgt å fokusere på tidlige refleksjoners innvirkning på opplevd avstand i den reproduserte musikken. Det er hovedsakelig tre årsaker til at jeg valgt å fokusere på avstand:

1. Fordi avstand et grunnleggende spatialt trekk
2. Fordi det er et begrep konsumenten kjenner fra før

### 3. Fordi det er relativt lett å kontrollere

Utgangspunktet for forsøket er følgende kunnskap og antagelser:

- En vet at nivåforholdet mellom direkte lyd og tidlige refleksjoner er en medvirkende faktor når høresansen skal bestemme avstanden til en lydkilde (Rumsey, 2001, s. 35). Begrepet direkte lyd blir beskrevet grundigere litt lenger nede, men er kort forklart den lyden som først når lytteren.
- I naturlige omgivelser er veien fra sanseapparat til bevisst erkjennelse av avstanden til en lydkilde, lett tilgjengelig for de fleste av oss (Bregman, 1990).
- En vet at man kan simulere økt avstand i reprodusert lyd ved å tilføre kunstige tidlige refleksjoner til en lydkilde (Rumsey, 2001).
- En vet at formidlingen av avstanden til lydkilder reprodusert i stereo, ikke er like forutsigbar som for lydkilder i naturlige omgivelser (Rumsey, 2001).
- En kan anta at veien fra sanseapparat til en bevisst erkjennelse av avstanden til en lydkilde i reprodusert lyd er relativt lett tilgjengelig for trente lyttere i optimale lyttesituasjoner.
- En vet at veien fra sanseapparat til bevisst erkjennelse er vanskeligere tilgjengelig for den utrente lytteren ved lytting til reprodusert lyd (Rumsey, Neher, Brookes, 2002).
- En vet at lytterrommets akustikk vil ha betydning for hørselens evne til å oppfatte reproduserte tidlige refleksjoner (Everest, 2001).
- Vi vet lite om hvordan ikke-optimale lyttesituasjoner påvirker avstandsopplevelsen i reprodusert musikk.

Det hele er kokt ned til et forsøk hvor utrente lyttere presenteres for innspilt musikk som er tilført tidlige refleksjoner i varierende grad. Hver oppgave i forsøket består av 4 versjoner av samme musikkseksempel, der forskjellen mellom versjonene er nivået på de tidlige refleksjonene. Forsøkspersonenes oppgave er å bedømme avstanden til lydkildene som reproduseres. Oppgavene utføres i to ikke-optimale lyttesituasjoner, samt en tilnærmet optimal. Forhåpentligvis vil forsøket gi oss litt mer kunnskap om konsumenters opplevelse av kunstige tidlige refleksjoner som er tilført musikken.

De ikke-optimale lyttesituasjonene representerer et enormt mangfold. Vi lytter i praksis til musikk overalt. Vi lytter til musikk i stua, på gata, i heisen, i bilen, på kjøpesenter og diskotek, for å nevne noe. Variasjonen mellom ulike lyttesituasjoner er stor, både med hensyn til akustikk, lydanlegg, støymengde og ikke minst lytterens eget fokus. For å gjøre det

praktiske forsøket håndterbart, har det derfor vært nødvendig å begrense det ytterligere. I forsøket har jeg valgt å fokusere på to lyttesituasjoner en kan anta oppstår relativt ofte innenfor hjemmets fire vegger. I tillegg har jeg tatt med en optimal lytteposisjon, som fungerer som en målestokk i forsøket. Den optimale lytteposisjonen heter på engelsk "sweet spot". Jeg har valgt å beholde det engelske navnet, da et norsk substitutt har vært vanskelig å finne. Forsøket er ytterligere noe begrenset med hensyn til forsøksmaterialets kompleksitet, men det kommer vi nærmere tilbake til litt senere.

### **1.2.1 Kommunikasjon med deltakeren**

Å beskrive klanglige opplevelser med ord er en utfordrende oppgave selv for den skolerte. Selv innen fagfelt som spesialiserer seg på kvalitative vurderinger av lyd, har man problemer med å komme fram til et felles begrepsapparat (Griesinger, 1999, s. 2). Når deltakerne i forsøket attpåtil er utrente i emnet, vil språket i seg selv representere en stor usikkerhetsfaktor. I stedet for verbal kommunikasjon har jeg derfor valgt å basere forsøket på en grafisk metode. Grafiske metoder har gjerne den fordel at de er intuitive å benytte for deltakeren.

I forsøket får deltakerne i oppgave å vurdere avstanden til lydkildene i den reproduserte musikken. Ved å bevege en skyveknapp på en graderingsskala, skal deltakeren forsøke å formidle sin opplevelse av avstanden til den enkelte lydkilde. Hver oppgave består, som nevnt, av 4 ulike versjoner av samme musikkeksempel. Spenningsmomentet ligger i hvorvidt deltakernes vurderinger gjenspeiler lydstyrken på de tidlige refleksjonene som er tilført musikken. En høyere lydstyrke på de tidlige refleksjonene, skal teoretisk innebære en økning av opplevd avstanden. Spørsmålet er i hvilken grad den utrente lytteren evner å oppleve det slik.

Det er ikke bare deltakernes iboende evner som avgjør hvorvidt de vil klare å identifisere ulikhetene mellom versjonene. Hørselens evne til å oppleve tidlige refleksjoner i reprodusert musikk avhenger blant annet av lydkildenes karakteristikk, lydbildets kompleksitet og ikke minst lyttesituasjonen.

### **1.2.2 Hypoteser**

På bakgrunn av de premissene som ligger i akustikken og psykoakustikken, ble det formulert fem hypoteser i forkant av forsøket. I hypotesene refereres det til tre ulike posisjoner. Posisjon 1 er i sweet spot, den tilnærmede optimale lytteposisjonen. Posisjon 2 er ved et bord i samme rom som høyttalerne, et stykke borte fra sweet spot. Posisjon 3 er i et tilstøtende rom. Det refereres også til de ulike versjonene av musikkeksemplene. Versjon 1 er alltid den versjonen

innenfor en oppgave hvor de tidlige refleksjonene er kraftigst, mens versjon 4 er den versjonen hvor refleksjonene er svakest. De fem hypotesene er som følger:

1. Feilprosenten vil være høyere ved posisjon 2 og 3 enn posisjon 1, og antageligvis høyere ved posisjon 3 enn posisjon 2.
2. Feilprosenten vil gå ned ved mer anvendt tid.
3. Det vil være høyere feilprosent mot lavere nivåer av tidlige refleksjoner.
4. Man vil se en utflating i graderingen mellom versjon 3 og 4.
5. Mer komplekst lydbilde og færre transienter vil gjenspeiles i større sprik i graderinger samt høyere feilprosent.

Hvorvidt de fem hypotesene blir bekreftet eller avkreftet, vil bli belyst i løpet av analysen senere i oppgaven.

### **1.3 Teoretisk fundament**

Det teoretiske grunnlaget for oppgaven faller i hovedsak inn under kategoriene lydteknikk, akustikk og psykoakustikk. I tillegg har litt kunnskap om digital signalprosessering vært viktig i tilnærmingen til deler av stoffet. Alle fagfeltene er selvsagt viktige og tett forbundet med hverandre, men siden det er søken etter menneskets persepsjon av romklang som er drivkraften bak arbeidet, havner psykoakustikken i en særstilling.

Da jeg startet arbeidet med denne hovedoppgaven på midten av nittitallet var det til dels vanskelig å finne egnet litteratur, spesielt innenfor psykoakustikken. Det har heldigvis endret seg litt. Det virker fortsatt slik at den jevne psykoakustiker plasserer romklang i en bakenforliggende strøm og lar den bli liggende der, men det finnes heldigvis unntak. David Griesinger synes å være den viktigste. Griesingers betydning for denne oppgaven kommer muligens ikke så godt fram i selve teksten, men hans allsidige artikler om både lydteknikk, akustikk og psykoakustikk danner et uvurderlig grunnlag for hele arbeidet. Han er en av de få som faktisk har tatt tak i den bakenforliggende strømmen romklangen gjerne plasseres i. Griesinger har blant annet definert tre ulike kategorier som beskriver vår persepsjon av romklangen. De tre kategoriene er gitt forkortelsene ESI, BSI og CSI som står for henholdsvis "Early Spatial Impression", "Background Spatial Impression" og "Continual Spatial Impression" (Griesinger, 1997). Jeg vil se litt nærmere på disse kategoriene i starten av kapittel 2.

Utover praktiske lærebøker og oppslagsverk innenfor emnet tenderer mye av det teoretiske materialet å være hentet fra forskere som søker svar på sine spørsmål gjennom praktiske forsøk. Francis Rumsey (m.fl.) er meget sentral, så også Barbara Shinn-Cunningham og Søren Bech.

Det er en gyldig innvending mot mitt arbeid at jeg i liten grad har tatt med kognitiv teori innen fagfeltet, på tross av at kognitive aspekter definitivt må spille en rolle i lytteforsøket. Opplevelsen av romklang og spatiale trekk i reprodusert lyd er et meget komplekst felt. Utelatelsen av det kognitive er først å fremst gjort i et forsøk på å begrense denne kompleksiteten.

Til slutt vil jeg trekke fram artikkelen ”Preference versus Reference: Listeners as participants in sound reproduction” av Durand R. Begault, som jeg kom over så sent som august 2006 (Begault, 2006). Det var en stor tilfredsstillelse og inspirasjon å endelig finne noen andre som har innsett behovet for å undersøke konsumenters lytteopplevelser.

## **1.4 Metode**

Det har vært relativt utfordrende å finne en god metode til å belyse konsumenters opplevelse av kunstig genererte tidlige refleksjoner. Arbeidet startet med en ide om å tilnærme seg problemet gjennom intervjuer, men det viste seg å være problematisk. Både forskningen og egen erfaring viser at opplevelse av spatiale trekk generelt, og romklang spesielt, er meget vanskelig å oversette til ord for den uskolerte. Det er sågar vanskelig for skolerte lyttere. I jakten på en metode gjennomførte jeg blant annet et forforsøk. Det var tre personer som deltok i forsøket, som var en kombinasjon av et intervju med lytteeksempler og et såkalt ABX-forsøk. En forklaring på hva et ABX-forsøk er, kommer litt lenger nede. Intervjuene var både meget interessante og lærerike, og trakk opp en lang rekke aspekter som med fordel kunne vært belyst av musikkvitenskapen. For eksempel, hvorfor assosierte ingen av forsøkspersonene i utgangpunktet kunstig romklang med rom? Men intervjuene bekreftet også til de grader en rekke utfordringer knyttet til omtalen av reproduserte lydbilder med uskolerte lyttere. Deltakerne utviste en meget god og kreativ språkbruk i omtalen av den reproduserte lyden. De brukte alle det som må sies å være gode metaforer til den klanglige opplevelsen, men utfordringen er at det ikke er noe enhetlig over den type beskrivelser. Hver person beskrev lyden på sin særegne måte. En annen utfordring er at lytterne i liten grad evnet å dissekere lydbildets bestanddeler på egenhånd, og her kommer det store problemet inn. I det man som forsker begynner å geleide intervjuobjektet i hennes analyse av lyden, starter man



samtidig en skoleringsprosess av både lytteevner og begrepsapparat som er problematisk for denne oppgavens problemstilling.

Et ABX-forsøk er en metode som har til hensikt å avdekke om en forsøksperson er i stand til å skille to lydeksempler fra hverandre. ABX-forsøk utføres ved hjelp av et dataprogram lagd for dette. Et eksempel på et slikt program er PCABX ([www.pcabx.com](http://www.pcabx.com)). I hver oppgave blir forsøkspersonen presentert for tre lydeksempler, A, B og X. Lydeksempel A og B er forberedt på forhånd. Lydeksempel X velges tilfeldig av dataprogrammet og tilsvarer enten A eller B. Forsøkspersonens oppgave er å forsøke å identifisere om X tilsvarer A eller B. Ved å repetere hver oppgave mange ganger, får man kartlagt om personen faktisk hører forskjell eller bare gjetter. Metoden er blant annet mye brukt for å avdekke kvalitative forskjeller mellom ulike lydkompresjonsformater, slik som mp3 og wma.

Metoden synes også å være anvendelig til å avdekke konsumenters evne til å oppfatte kvalitative forskjeller knyttet til bruken av kunstig romklang, være seg både terskelverdier og parameterendringer. Resultatene fra forforsøket var interessante, og det syntes derfor meget attraktivt å velge en metode som gir så statistisk gode svar. Metoden ble allikevel valgt bort til fordel for en grafisk metode som synes å være enda bedre egnet.

Grafiske metoder for synliggjøring av subjektive opplevelser (*engelsk: Graphical Elicitation Techniques*) synes å være velegnede i fagfeltet vi beveger oss i. Den åpenbare grunnen, som også allerede er nevnt, ligger i den delvise løsrivelsen fra språket. Helt fri fra språket blir man ikke, men grafiske metoder gir i alle fall lytteren mulighet til å uttrykke sin opplevelse uten å måtte gå via språket. Den grunnleggende tanken er at opplevelsen av spatiale trekk i reprodusert lyd er tettere knyttet til det visuelle enn den er til ord. Man kan anta at et nære slektskap mellom uttrykket (lyden) og metoden vil innebære at behovet for tolkning fra respondentens side blir lavere (Rumsey, Bruyn, Ford, 2001, s.2). Det nære slektskapet ligger i stor grad i at spatiale trekk i et lydbilde lett kan knyttes til fysiske attributter ved lydkilden(e) og deres omgivelser. Opplevelsen av fysiske attributter på sin side, er gjerne lettere å formidle grafisk enn med ord. Det gamle ordtaket ”et bilde sier mer enn tusen ord” må sies å ha en viss relevans.

I praksis benytter man gjerne tegning på frihånd eller datagenererte grensesnitt for å fange opp det grafiske uttrykket. Tegning på frihånd utføres med fordel ved hjelp av penn og papir. I de datagenererte grensesnittene benytter man gjerne skyvbare knapper på ulike graderingsskalaer til å formidle det man opplever. En vesensforskjell mellom de to tilnærmingene er at tegning på frihånd generelt åpner for en langt større variasjon mellom svarene fra de ulike deltakerne. Det kan i utgangspunktet synes lite ønskelig, men man kan

argumentere at det enkelte uttrykk da vil være nærmere lytterens egen opplevelse og at utfordringen flyttes over på forskeren som skal analysere besvarelsene. De datagenererte grensesnittene vil på sin side gi et mer unisont uttrykk de ulike deltakerne i mellom, men man øker samtidig usikkerheten om hvorvidt deltakerne har vurdert det man ønsker de skal vurdere, eller om de har vurdert helt andre aspekter.

En rekke faktorer vil påvirkes av kommunikasjonen mellom forskeren og deltakeren i forkant av forsøket, implisitt de føringer som formidles til deltakeren. Man forholder seg her til den klassiske balansegangen mellom ønsket og uønsket påvirkning av deltakeren. Mer informasjon til deltakerne vil sannsynligvis gi mer entydige svar, men man vil samtidig øke faren for at forskerens forventninger påvirker deltakerens besvarelse.

Jeg har valgt å benytte et datagenerert grensesnitt med skyvbare knapper i mitt lytteforsøk. Det er i hovedsak to aspekter som begrunner valget av grensesnitt. For det første fokuserer forsøket på opplevelsen av en enkel dimensjon, dybdedimensjonen. Koplingen mellom opplevelse og graderingsskalaens ene variable dimensjon, vil antakeligvis være meget intuitiv for deltakeren. Dette senker muligheten for feiltolkning fra deltakerens side, og dermed også muligheten for feiltolkning i analysen. Det andre aspektet ligger i oppgavens utforming. Deltakerne blir bedt om å sammenligne fire lydeksempler innenfor hver enkelt oppgave. En kan anta at deltakeren vil veksle fram og tilbake mellom lydeksempelene, og samtidig justere graderingen av det enkelte eksempel. Et datagenerert grensesnitt synes meget velegnet til den type sammenligning.

Utgangspunktet for den grafiske metoden jeg benytter, er i stor grad hentet fra vitenskapen om lyd kvalitet i reprodusert lyd. Mitt arbeid kan på flere vis plasseres inn i denne tradisjonen. De grunnleggende premisene i akustikken, psykoakustikken og lydteknikken er akkurat de samme. Samtidig skiller mitt fokus seg fra tradisjonen på et par grunnleggende punkter. Det viktigste skillet ligger i mitt fokus på den uskolerte lytteren og at det er dennes opplevelse som er i fokus. Mitt fokus står i sterk kontrast til vitenskapens sterke fokus på skolering av forsøkspersonene i kritisk lytting i forkant av lytteforsøk (Rumsey, Neher, Brookes, 2002). Det primære målet for vitenskapen rundt lyd kvalitet er å øke kunnskapen om premisene som ligger til grunn for lytterens opplevelse, slik at man bedre kan kontrollere disse. Til dette arbeidet er det en fordel desto bedre forsøkspersonen er i stand til å dissekere de spatiale trekkene i et lydbilde. Hensikten med min problemstilling derimot, kan beskrives som et ønske om å finne ut hvor bra konsumenter klarer å dissekere et lydbilde i utgangspunktet. Min tilnærming til metoden er i så måte noe eksperimentell. Man må anta at

en konsekvens av mitt fokus på uskolerte lyttere vil innebære et større sprik i besvarelsene, enn om man hadde benyttet skolerte forsøkspersoner.

Det er berettiget og relevant å spørre om metoden er anvendelig på uskolerte lyttere i det hele tatt. Kan man hente ut noe som helst sikkert fra besvarelsene til uskolerte lyttere? Spørsmålet om anvendelighet er for så vidt et relevant spørsmål til hele vitenskapen, uavhengig av skolerte eller uskolerte forsøkspersoner. Forutsatt at man aksepterer at metoden er anvendelig på skolerte lyttere, tror jeg svaret avhenger mye av hvilke spatiale trekk man ber den uskolerte lytteren om å fokusere på. En kan anta at jo mer kjent det spatiale trekket er for lytteren i utgangspunktet, desto bedre vil metoden fungere. Det er nettopp av denne grunn jeg har valgt å fokusere på opplevelsen av avstand. Man kan anta at avstand er et så kjent begrep for allmennheten, at deltakerne enkelt vil forstå hvilken oppgave de er ment å utføre og hva de skal fokusere på for å utføre den.

Det er varierende i hvilken grad spatiale trekk i reprodusert lyd lar seg kvantifisere. Grovt generalisert kan man si at breddedimensjonen i større grad lar seg kvantifisere enn dybdedimensjonen. Avstand er i så måte meget vanskelig å kvantifisere. Graderingsskalaen som benyttes i lytteforsøket inneholder derfor ingen form for kvantifisering.

Det noe ambisiøse fokuset på konsumenter og ikke-optimale lyttesituasjoner, introduserer flere usikkerhetsfaktorer som må ligge til grunn for analysen i etterkant av forsøket. Noen faktorer er både åpenbare og gode ankepunkter mot metoden. Det største ankepunktet er antakeligvis at det er vanskelig å vite om alle forsøkspersonene fokuserer på det man ønsker de skal fokusere på. Usikkerhetsfaktorene kommer jeg selvsagt tilbake til i analysen.

## **1.5 Forutsetninger hos leseren**

En hovedoppgave skal helst formidle forfatterens grunnleggende kunnskap om emnet, og innebærer derfor gjerne beskrivelser av en rekke grunnprinsipper for vitenskapsfeltet. Den obligatoriske beskrivelsen av grunnprinsippene fungerer også som hjelp til leseren som ikke kjenner fagfeltet. Slik er det også i denne oppgaven. Enkelte ting blir grundig beskrevet, men det er umulig å dekke alt. Enkelte ting må man anta at leseren kjenner til på forhånd, eller i det minste vil ha evne til å forstå ut fra sammenhengen, uten detaljert kunnskap på forhånd.

Jeg har valgt å vie min grundige beskrivelse spesifikt til de tidlige refleksjonene, og aspekter direkte knyttet til disse, være seg akustiske, psykoakustiske eller lydtekniske. Det innebærer at jeg forutsetter at leseren har en grunnleggende forståelse av alle de tre feltene, men kanskje spesielt psykoakustikken. Noen sentrale begreper blir definert i

begrepsavklaringen like nedenfor. Andre vil bli forklart etter hver som de dukker opp, men uansett plassering vil definisjonene være kortfattede. Kapittel 2 er nok det kapittelet som er vanskeligst tilgjengelig for leseren som ikke er inne i faget. De resterende kapitlene bør være litt lettere tilgjengelig.

## 1.6 Avgrensing

For å gjøre dette temaet håndterbart, har det vært nødvendig med et snevert fokus. Et snevert fokus har absolutt sine fordeler. Det er antakeligvis da man har størst mulighet til å avdekke noe vesentlig, men det innebærer også at viktige ting må overses. Leser man problemstillingen er den åpenbare taperen de sene refleksjonene, men det er et valg som kan forsvares ut fra både akustiske, psykoakustiske og lydtekniske kriterier. Det er spesielt i omgangen med konsumenters lyttesituasjoner jeg har måttet overse en rekke ting.

Jeg overser helt reproduksjon i hodetelefoner, dette på tross av at det har vært en markant økning i bruken av mp3-spillere i senere tid (Vaage, 2005, s. 36). Mp3-spillere benyttes gjerne sammen med hodetelefoner. Opplevelsen av spatiale trekk i hodetelefoner fortjener absolutt sitt fokus, men lytting i hodetelefoner skiller seg markant fra annen lytting på flere områder. De to mest opplagte forskjellene mellom hodetelefoner og andre høyttalere er lytterrommets manglende innvirkning på førstnevnte, samt total kanalseparasjon mellom ørene. Hodetelefoner krever dermed et noe annet fokus og er dermed utelatt.

I lytteforsøket har jeg kun valgt ut to lyttesituasjoner som er karakteristiske for konsumenters musikklytting. Det betyr at jeg overser mye annet. Vi lytter til musikk i et enormt mangfold av steder. Mye av det som skrives i oppgaven vil være gyldig for flere av disse stedene, men deres særpreg behandles ikke separat. Listen kan bli meget lang hvis jeg skal nevne alt ved konsumenters lyttesituasjon som ikke berøres. Et fellestrekk for en del, knytter seg til ulike former for degradering av det innspilte lydsignalet. Være seg på grunn av kompresjonsalgoritmer, variasjoner i avspillingsanlegg eller støyende omgivelser.

Den optimale lyttesituasjonen blir i liten grad utdypet. En grundig definisjon av den optimale lyttesituasjonen ville vært stoff til en hel oppgave i seg selv. Det må sies at selve begrepet "optimal lyttesituasjon" er en tenkt størrelse. Hele lyd- og hifi-verden jakter på denne situasjonen, og den er ikke lett å finne. Min bruk av begrepet må dermed leses som "det nærmeste man kommer en optimal lyttesituasjon".

Til slutt må det nevnes at jeg helt overser selve musikken i oppgaven. Både den musikalske helheten og enkeltinstrumentenes særpreg behandles i liten grad. Mitt fokus er litt

til siden for selve musikken, men arbeidet utføres med visshet om at det musikalske kan ha stor betydning for konsumentens opplevelse av spatiale trekk.

## 1.7 Begrepsavklaring

### *Direkte lyd*

Som navnet innebærer er den direkte lyden de lydbølgene som beveger seg direkte fra lydkilde til lytter. I motsetning til reflektert lyd, som på sin vei har vært innom en eller flere overflater før den når lytteren. Den direkte lyden vil være de lydbølgene som har reist korteste veien mellom lydkilden og lytteren, og er dermed de lydbølgene som treffer lytteren først. Forutsatt at det er fri sikt mellom lydkilde og lytter, vil lyden ha beveget seg i en rett linje. Den direkte lyden spiller en avgjørende rolle i hørselens kartlegging av de ulike lydkildene vi hører. Den direkte lyden er også sentral i opplevelsen av romklang og spesielt de tidlige refleksjonene, da vår opplevelse av romklangen i stor grad styres av dens relasjon til den direkte lyden.

### *Romklang (tidlige og sene refleksjoner)*

Romklang er lydbølger som har vært reflektert fra en eller flere overflater på sin vei mellom lydkilde og lytter. Romklang oppstår overalt i naturen, men den karakteristiske opplevelsen av romklang forbindes hovedsakelig med lukkede rom. Akustikken deler romklangen inn i tidlige- og sene refleksjoner. De tidlige refleksjonene er de første refleksjonene som treffer lytteren etter den direkte lyden. De tidlige refleksjonene har vanligvis bare vært innom noen få overflater før de treffer lytteren. De karakteriseres ved å være få i antall og er relativt distinkte, sammenlignet med de sene refleksjonene.

Når en lydkilde genererer lyd i et rom vil det oppstå flere og flere refleksjoner etter hvert som lydbølgene beveger seg mellom de ulike overflatene i rommet. Mens de tidlige refleksjonene er relativt få i antall, består de sene refleksjonene av et enormt antall refleksjoner. Mange vil nok kople den distinkte opplevelsen av romklang til de sene refleksjonene, spesielt i omgangen med konsertsaler, kirker og kunstige romklangsimulatorer. Dette skyldes at de tidlige refleksjonene gjerne oppfattes som en del av den direkte lyden, mens de sene refleksjonene er den klanglige halen som følger etter lyden.

### *Kunstig romklang*

Betydningen av begrepet "kunstig romklang" kan synes å være selvforklarende. Den enkle definisjonen er at romklangen er kunstig skapt, i motsetning til den romklang som oppstår i naturlige omgivelser. I kapittel 3 vil vi se at begrepet er noe mer komplisert. En definisjon

som tar høyde for komplikasjonene er som følger: *Kunstig romklang er romklang som er initiert ved reproduksjon av lydkilden, eller som manipuleres uavhengig av den direkte lyden.*

#### *Algoritme (romklungalgoritme)*

Kunstig romklang skapes gjerne ved hjelp av digitale romklangsulatorer. Algoritme og romklungalgoritme er begreper som refererer til oppbyggingen av programmene i romklangsulatorer. En romklangsulator kan for eksempel ha en algoritme som heter konsertsal, en som heter kirke og en annen som heter badrom.

#### *Predelay x 2*

Predelay (*norsk: initial forsinkelse*) er en parameter som i utgangspunktet benyttes til å beskrive forsinkelsen i ankomsttidspunktet til den første tidlige refleksjonen, hvor forsinkelsen måles ut fra ankomsttidspunktet til den direkte lyden. Noen produsenter av kunstige romklangsulatorer benytter også ordet til å beskrive forsinkelsen i ankomsttidspunktet til de sene refleksjonene. Romklangprogrammet TrueVerb som ble benyttet i denne oppgavens lytteforsøk er en av disse.

#### *Innstikk (plugin)*

Innstikk er bedre kjent under det engelske begrepet plugin. Innstikk er små dataprogrammer som koples på et større dataprogram for å utføre spesifikke oppgaver. Et eksempel er det store dataprogrammet ProTools og innstikket TrueVerb. ProTools er et program for å ta opp, redigere og mikse lyd med mange spor/kanaler. TrueVerb er et innstikk som simulerer romklang. Innstikket kan koples til en eller flere av kanalene i ProTools.

#### *Ekkofritt rom*

Et ekkofritt rom er et rom hvor lyden ikke reflekteres fra rommets overflater. Ekkofrie rom er spesialbygde rom hvor veggene består av tykke lag med lydabsorberende materialer. Ekkofrie rom er med andre ord, rom uten romklang.

#### *Fritt felt*

I omtalen av lydbølgers utbredelse tar man gjerne utgangspunkt i lydens bevegelse i det som kalles fritt felt. Fritt felt er omgivelser hvor lyden på sin ferd ikke reflekteres fra noen overflater. Det er stort sett i fritt fall noen kilometer over bakken og i ekkofrie rom man kan oppleve det. Alle andre steder vil lyden reflekteres i større eller mindre grad. Lydens utbredelse i fritt felt danner allikevel grunnlaget for omtalen av lyd, og kunnskapen er gyldig i en rekke sammenhenger. Dersom den direkte lyden beveger seg uten hindringer mellom lydkilde og lytter, vil den kunne betraktes ut fra samme prinsipper som lyd i fritt felt. Et av

grunnprinsippene for lydens bevegelse i fritt felt er det som på engelsk heter ”inverse square law”. På norsk blir det noe slikt som ”loven om det omvendte kvadrat” (*egen oversettelse*). Loven innebærer at for hver gang man dobler avstanden til lydkilden, blir lydtrykket redusert til en fjerdedel av hva det var. Overført til desibel (dB) vil en dobling av avstanden til lydkilden innebære en reduksjon av lydstyrken på 6 dB.

#### *Desibel*

I følge ordboken er desibel (forkortes: dB) en ”logaritmisk måleenhet for lydstyrke, støy” (”Bokmålsordboka”, 2006). Desibelskalaen er i for seg ikke forbeholdt lyd, men det er her den gjerne benyttes. I den vanligste bruken av desibel refererer verdien 0 dB til det laveste lydtrykket menneskets hørsel er i stand til å oppfatte. I teknisk omtale av lyd benyttes desibelskalaen for å synliggjøre ulikheter mellom ulike lydnivåer. I forberedelsene til oppgavens lytteforsøk ble lyden tilført tidlige refleksjoner med ulik lydstyrke. I omtalen av de ulike nivåene som ble satt, er en versjon/lydstyrke valgt som referansepunkt. Referansepunktet omtales som 0 dB. De andre versjonene/lydstyrkene beskrives i relasjon til referanseverdien med pluss- og minusverdier.

#### *Hertz og kilohertz*

Hertz (forkortes: Hz) er antageligvis unødvendig å definere, men for ordens skyld er hertz en enhet for frekvens, som beskriver antall svingninger per sekund. 1 kilohertz (forkortes: kHz) er 1000 svingninger per sekund.

#### *Lydstyrke og opplevd lydstyrke (loudness)*

Begrepet lydstyrke retter seg inn mot to separate områder. På den ene siden har men den fysiske målbare lydstyrken. På den andre har man den opplevde lydstyrken (*engelsk: loudness*). På steder der ordet lydstyrke benyttes alene, refererer dette til den målbare lydstyrken. Desibelskalaens mål av lydstyrke tar utgangspunkt i lydbølgenes bevegelse gjennom luft og gjenspeiler således det fysiske lydtrykket i et lydsignal. Menneskets opplevelse av lydstyrken til et lydsignal er mer kompleks og mindre forutsigbar. Vår opplevelse av lydstyrke påvirkes i stor grad av frekvensrikdommen i et lydsignal, i motsetning til fysiske målinger av lydtrykk, som ikke gjør det.

#### *Filtrering (EQ)*

Filtrering er et lydteknisk begrep som innebærer frekvensspesifikk forsterking eller reduksjon av lyd. Filtrering er bedre kjent under det engelske begrepet equalizing. Forkortelsen ”EQ” er hentet fra det engelske begrepet.

### *Kompresjon*

Kompresjon utføres av kompressorer og innebærer automatisk nivåjustering av lyd. Kompressorer er et lydteknisk verktøy som benyttes i lydstudioer for å jevne ut lydnivået.

### *De-esser*

En de-esser er en kompressor som komprimerer et avgrenset frekvensområde i lyden. Verktøyet brukes mest til å senke lyden av markante s'er ved miksing av vokal. I forbindelse med denne oppgaven ble det benyttet til å senke lyden av undertegnede fingre som gled opp og ned strengene på en gitar.

### *Insert*

Insert er et engelsk begrep som refererer til et sted i signalgangen i en miksepult hvor man gjerne kopler inn ulike effekter, slik som kompressor, de-esser og equalizer.

### *Mikrofoners karakteristikk (omni, nyre og 8-tall)*

Mikrofoner konstrueres med ulike retningskarakteristikk. Retningskarakteristikken forteller om hvor retningsbestemt en mikrofon vil fange opp lyd. En mikrofon med omnikarakteristikk fanger inn lyden like kraftig fra alle retninger. Mikrofoner med nyrekarakteristikk er mer retningsbestemte. Lyd som treffer på midten, framme på mikrofonen formidles kraftigst, mens lyd som kommer inn bakfra og/eller til siden fanges inn i mindre grad. Som det visuelle utseende til tallet 8 tilsier, formidler en mikrofon med 8-tallskarakteristikk sterkeste lydbølger som kommer inn fra to sider av mikrofonen.

### 8-tall

### *Nærfelt, kritisk distanse (romradius) og fjernfelt*

I lukkede rom vil lyden som kjent reflekteres fra overflatene i rommet, med den konsekvens at loven om det omvendte kvadrat ikke lenger gjelder. I lukkede rom opererer man derfor med tre begreper, nærfelt, kritisk distanse og fjernfelt. I nærfeltet vil lydtrykket fra den direkte lyden være kraftigere enn romklngen i rommet. Beveger man seg bort fra lydkilden, kommer man etter hvert til punktet som kalles kritisk distanse. Lydstyrken til den direkte lyden følger som nevnt loven om det omvendte kvadrat, og en bevegelse bort fra lydkilden vil dermed innebære en gradvis senking av den direkte lyden. Romklngen i et lukket rom følger ikke denne loven. Lydtrykket er i stedet relativt konstant. Ved den kritiske distansen, som også kalles romradius, er lydtrykket til den direkte lyden og romklngen likt. Beveger vi oss videre kommer vi inn i fjernfeltet. Her forblir lydstyrken i rommet konstant, mens den direkte lyden



fortsetter å reduseres. Jo lenger man kommer inn i fjernfeltet, desto dårligere blir blant annet forståeligheten (*engelsk: intelligibility*).

### *Maskering*

Maskering er et begrep som benyttes i psykoakustikken og betydningen er relativt selvforklarende. Maskering betyr i enkelhet at en svakere lyd ikke oppfattes av hørselen på grunn av en kraftigere lyd. Prinsippene for maskering er langt fra like enkle, men kort forklart vil graden av maskering avgjøres av et samspill mellom lydstyrke og frekvenskompleksitet. Maskeringen kan deles inn i tre ulike grupper. Den første er den samtidige maskeringen, hvor lyden som maskerer og lyden som blir maskert eksisterer på samme tid. Den andre er maskering som oppstår der en kraftig lyd maskerer en etterfølgende svakere lyd. Disse to er de absolutt vanligste. Den siste maskeringseffekten, som ikke oppstår fullt så ofte, er maskering forover (*engelsk: forward masking*). Maskering forover innebærer at en svakere lyd enkelte ganger kan maskeres av en sterkere etterfølgende lyd.

Maskering er for øvrig tett forbundet med et annet psykoakustisk begrep som heter fusjon (*engelsk: fusion*). Fusjon er sammensmeltingen av lyder, og er meget viktig i vår opplevelse av musikk, men behandles ikke i denne oppgaven.

### *Spatial avmaskering*

Spatial avmaskering (*engelsk: spatial unmasking*) er en psykoakustisk effekt som kan oppstå når man øker avstanden mellom den maskerende lydkilden og den maskerte lydkilden. I praktisk tilnærming innebærer det blant annet at en lytter vil ha lettere for å oppfatte og forstå en lydkilde dersom avstanden mellom lydkilden og en forstyrrende lydkilde økes.

Det motsatte av spatial avmaskering er spatial maskering.

### *Omhylling*

Omhylling (*engelsk: envelopment*) er en opplevelse som er spesielt ettertraktet i konsertsaler, men er også aktuell i opplevelsen av reprodusert lyd. Omhylling er opplevelsen av å være omkranset av lyden som skapes, at man tar del i den og ikke opplever den som noe bortenfor en selv.

### *Eksternalisering*

Eksternalisering (*engelsk: externalization*) er det motsatte av omhylling. Man befinner seg ikke i lyden, men opplevelsen er som å betrakte lyden på avstand. Konsumenters omgang med reprodusert musikk i høyttalere vil normalt innebære en opplevelse av eksternalisering, ikke omhylling.

### *Etterklangstid (RT60)*

Romklangens etterklangstid er enkelt forklart en tidsangivelse som sier hvor lang tid det tar før romklangen i et rom dør ut etter at lydkilden har sluttet å generere lyd. Konsertsaler og kirker har gjerne en lang etterklangstid, mens små kompakte rom vanligvis har en kort etterklangstid. Akustikkens måler etterklangstiden ved å se på tiden det tar fra refleksjonene er på sitt høyeste til lydstyrken har falt med 60 dB. Dette målet benevnes med RT60, hvor RT står for det engelske "reverberation time". I mindre rom benytter man gjerne både RT30, men RT60 er langt på vei det vanligste. Konsertsaler har gjerne en etterklangstid fra 1 sekund og oppover, mens stuer og soverom gjerne ligger på noen hundre millisekunder (ms).

### *Diffusjon*

Diffusjon er et begrep som omtaler graden av oppbryting/spredning en lydbølge utsettes for når den reflekteres fra en overflate. En stor, glatt overflate vil spre lydenergien i mindre grad enn en oppbrutt overflate. I beskrivelsen av lyd sammenligner man gjerne lydbølger med stråler. En stor, glatt overflate med lav grad av diffusjon vil reflektere en innkommende lydbølge (stråle) i en enkelt utgående stråle. En overflate med høy diffusjon vil fordele energien fra den innkommende strålen i mange utgående stråler. De utgående strålene spres i ulike retninger.

### *Transient*

Transienter er de uregelmessige lydsvingninger som oppstår i startfasen av lydhendelser. De etterfølges gjerne av at lyden går over i mer regelmessige svingninger. Transienter er tett forbundet med anslag på instrumenter og konsonanter i tale/sang. En skarptromme er et eksempel på en lydkilde som er rik på transienter, mens en fløyte gjerne genererer transienter i svakere grad. Transientene er gjerne rike på høyere frekvenser, og har generelt noe høyere lydenergi enn de regelmessige svingningene. Transienter frambringer derfor gjerne tydeligere refleksjoner (romklang).

### *Den optimale lytteposisjonen og sweet spot*

Det ble nevnt ovenfor at den optimale lyttesituasjonen er en tenkt størrelse, og at min bruk av begrepet derfor må leses som "det nærmeste man kommer en optimal lyttesituasjon". Det er en rekke årsaker til at den optimale lyttesituasjonen er vanskelig å konstruere. Mange av årsakene kan knyttes til konstruksjonen av komponentene i stereoanlegget som skal produsere lyden, samt interaksjonen mellom stereoanlegg og lytterrommet i overføringen til lytteren. Den optimale lyttesituasjonen kompliseres ytterligere av aspekter knyttet til selve

kildematerialet, altså innspillingen som skal reproduseres. Den gjengse oppfatningen er at ulikt kildemateriale stiller ulike krav til reproduksjonen.

Min pragmatiske tilnærming til begrepet gjør bildet noe enklere, men det er fremdeles en rekke faktorer som definerer den optimale lyttesituasjonen ved lytting til reprodusert musikk i stereo. Faktorene kan grovt sett deles inn i tre kategorier. Kategoriene er avspillingsmedium, avspillingsrom og lytteren. Når det gjelder avspillingsmedium vil jeg nøye meg med å nevne noen grunnleggende premisser som må være tilfredsstillt. Disse er:

- Kildematerialet må være av høy oppløsning med hensyn til lyd kvalitet.
- Stereoanlegget må reprodusere lyd i full båndbredde, altså i hele det hørbare frekvensområdet. Anlegget må også ha en flat frekvensrespons. At frekvensresponsen er flat, betyr at alle frekvenser blir reprodusert like kraftig.
- Stereoanlegget må reprodusere lyden med lav grad av forvrengning og støy.
- Stereoanlegget må ha god transientrespons.
- Høytalernes posisjon må ha en god, symmetrisk plassering i lytterrommet.

En optimal lyttesituasjon knytter også en rekke utfordringer til lytterrommets utforming. Dette vil bli behandlet grundig i kapittel 3. Kravene til godt kvalifiserte lytterrom er så omfattende med hensyn til utforming og innredning, at man ikke finner slike rom hjemme hos konsumenten.

Når det gjelder lytteren er det hovedsakelig lytterens posisjon i forhold til høyttalerne som betinges av det optimale. Ved en optimal lyttesituasjon er lytterens posisjon begrenset til det lille punktet midt mellom, foran høyttalerne som på engelsk kalles "sweet spot". Man finner punktet ved å konstruere en likebeint trekant med høyttalere og lytteposisjon i hvert av hjørnene. Lytteren skal befinne seg like bak spissen av trekanten. Linjen mellom høyttalerne danner grunnlinjen og vinkelen fra lytteposisjon til høyttalerne skal være  $\pm 30^\circ$  (Rumsey, 2001, s. 53). Størrelsen og posisjonen for sweet spot kan variere litt på bakgrunn av enkelte andre faktorer, men det er relativt snevert og det er altså det eneste stedet man kan oppleve en presis gjengivelse av stereolydens spatiale trekk. Jo mer man beveger seg bort fra sweet spot, desto mer fordreies og etter hvert innsnevres normalt lydbildet.

## 2 Tidlige refleksjoner

Tidlige refleksjoner er en viktig bestanddel i romklang. De har stor betydning for lyden vi hører. Fra starten av en produksjon og fram til konsumentens ører vil lyden tilføres tidlige refleksjoner flere steder i prosessen. Noen steder vil påvirkningen være minimal, gjerne uhørbar, men ofte vil refleksjonene farge lyden markant, i både positiv og negativ retning. Behersking og kontroll av tidlige refleksjoner er en veldig viktig faktor i profesjonelle musikkproduksjoner. En produsent bør naturlig nok være bevisst de refleksjoner han fanger opp med mikrofoner og tilfører ved effektprosessering, men viktige er også refleksjonene i rommet der produsenten gjør sine valg.

I dette kapitlet skal jeg gå litt i dybden på tidlige refleksjoner og deres innvirkning på lyden vi hører. For å plassere de tidlige refleksjonene inn i et mer helhetlig perspektiv, vil jeg starte med et kort blikk på noen psykoakustiske rammebetingelser som ligger til grunn for vår opplevelse av spatiale trekk som følge av romklang generelt. Resten av kapitlet vil fokusere på de tidlige refleksjonene. Først fra et akustisk perspektiv, for deretter å bevege meg over mot psykoakustikken igjen. I løpet av kapitlet vil jeg trekke fram noe av den interessante og viktige forskningen som er gjort på tidlige refleksjoner, og håper på den måten å få fram noe av bredden rundt dette temaet.

### 2.1 Hvordan opplever vi romklang

Det er få forskere innen psykoakustikken som har tatt tak i vår opplevelse av romklang. David Griesinger er faktisk den eneste undertegnede har kommet over som ser på basale trekk i persepsjonen av romklang, og hvordan den tar del i den lyden vi opplever. Denne delen baserer seg derfor i helhet på hans teorier.

Griesinger har definert tre kategorier som på forskjellig vis kopler romklangen til høresansens basale oppdeling av lyden i strømmer (*engelsk: streams*). Strømmer er kjent fra Bregman (Bregman, 1990). De tre kategoriene er:

1. CSI som står for continuous spatial impression. Norsk: vedvarende spatiale inntrykk (*egen oversettelse*)
2. BSI som står for background spatial impression. Norsk: spatiale inntrykk av bakgrunnen (*egen oversettelse*)
3. ESI som står for early spatial impression. Norsk: tidlig spatiale inntrykk (*egen oversettelse*)

Vedvarende spatiale inntrykk (CSI) kan bare oppstå når en lydhendelse er vedvarende (Griesinger, 1997, s. 7). Når den direkte lyden fra en lydkilde er vedvarende og ikke kan deles opp i separate hendelser, kan interaksjonen med reflektert energi (romklangen) gi en opplevelse av full omhylling eller romlighet (*engelsk: spaciousness*) som synes å være koplet til lydkilden (Rumsey, 2001, s. 44). Opplevelsen av CSI forutsetter vedvarende lydhendelser, i motsetning til ESI og BSI som oppstår ved diskrete lydhendelser. Vedvarende lydhendelser oppstår sjelden i musikken, hvor man i stor grad opererer med stadig skiftende lydkilder. CSI er dermed lite involvert i opplevelsen av musikk.

Spatiale inntrykk av bakgrunnen (BSI) er opplevelsen man mest forbinder med romklang, eller med andre ord, klangen av rommet. BSI knyttes dermed til romklangens sene refleksjoner. BSI oppstår når det er mulig for hørselen å dele den innkommende lyden i en forgrunn- og en bakgrunnsstrøm (Griesinger, 1997, s. 11). Opplevelsen av BSI oppstår på bakgrunn av de samme fysiske attributtene som CSI, men i motsetning til CSI, oppstår BSI når den direkte lyden er fraværende. Med andre ord, og litt forenklet, oppstår BSI der det er pauser i musikken. På grunn av den korte etterklangstiden i mindre rom er BSI en opplevelse som hovedsakelig oppstår i større rom. Opplevelsen av BSI er viktig i opplevelsen av innspilt romklang, men vil være mindre involvert i interaksjonen mellom den reproduerte lyden og lytterrommet til den konsumenten. Ved lytting til et orkester i en konsertsal vil opplevelsen av BSI være større desto høyere lydstyrken er (Griesinger, 1999, s. 7).

Tidlig spatiale inntrykk (ESI) knyttes til refleksjoner som ankommer i løpet av en lydhendelse og inntil 50 millisekunder etter lydhendelsen har stanset. ESI er dermed tett forbundet med tidlige refleksjoner, som er fokus for denne oppgaven. ESI oppleves i stor grad tilknyttet lydkilden, ikke som separate lydhendelser. ESI bidrar i så måte til en rekke spatiale

trekk forbundet med selve lydilden. ESI er også grunnlag for opplevelsen vi har av romlighet i små rom.

## **2.2 Akustikken**

### **2.2.1 Hva er tidlige refleksjoner**

Det er en viss inkonsekvens i omtalen av tidlige refleksjoner og romklang. De aller fleste vil nok definere tidlige refleksjoner som en del av romklangen, men det er ikke uvanlig å omtale de tidlige refleksjonene som noe eget mens man forbeholder ordet romklang til de sene refleksjonene. Dette er kanskje særlig tydelig innenfor musikkteknologien (Lexicon, 2001, s.5-3), men det gjelder ikke bare der. De færreste vil nok ta på seg oppgaven med å argumentere at tidlige refleksjoner ikke er romklang, både med hensyn til refleksjonenes tilblivelse og ikke minst informasjonen hørselen henter ut av dem. Ulikhetene i omtale skyldes nok heller ingen underliggende uenighet, men gjenspeiler nok mer pragmatiske og pedagogiske valg med hensyn til innfallsvinkel og formidling til mottakeren. Uansett omtale er det viktig å understreke at tidlige refleksjoner er en viktig bestanddel i det vi kaller romklang.

Vitenskapen rundt romklang i musikalsk sammenheng tar normalt utgangspunkt i lyd som oppstår i lukkede rom. I tillegg forutsetter man gjerne at lyden kan bevege seg uhindret i en rett linje fra lydkilde til lytter. Slik er det også i denne beskrivelsen.

Når lyd sendes ut fra en lydkilde i et rom vil det første vi hører være de lydbølgene som har reist den korteste banen mellom oss og lydkilden. Dette er den direkte lyden. Tidlige refleksjoner er de første refleksjonene som treffer oss etter den direkte lyden. Refleksjonene stammer fra nærliggende overflater. Antallet refleksjoner som ankommer lytteren øker gradvis. Etter hvert er antallet så stort at man er over i de sene refleksjonene.

Det har vært vanskelig å finne en eksakt definisjon på tidlige refleksjoner, eller retttere sagt, skillelinjen mellom tidlige og sene refleksjoner. Noen steder hevdes det at kun refleksjoner av første orden, altså de som bare har vært innom en enkelt overflate før de når lytteren, kvalifiserer til å være tidlige (Kapralos, 2003, s. 24). En slik definisjon blir for streng, men det er interessant å merke seg at man selv i enkelte profesjonelle romklangsimulatorer generer alle tidlige refleksjoner som om de var av første orden (Geoffrey, 2001, s. 106). Generelt benevner man refleksjoner som ankommer innen 50-80 ms etter den direkte lyden som tidlige refleksjoner. Teoretisk sett kan tidlige refleksjoner være av hvilken orden som helst, men vil i de fleste sammenhenger være av lavere orden. Den eksakte skillelinjen mellom tidlige og sene refleksjoner er ikke veldig interessant i denne

sammenheng. Det viktige her er å se hvordan de første refleksjonene som når lytteren påvirker lyden og vår opplevelse av den.

### **2.2.2 En enkelt refleksjon**

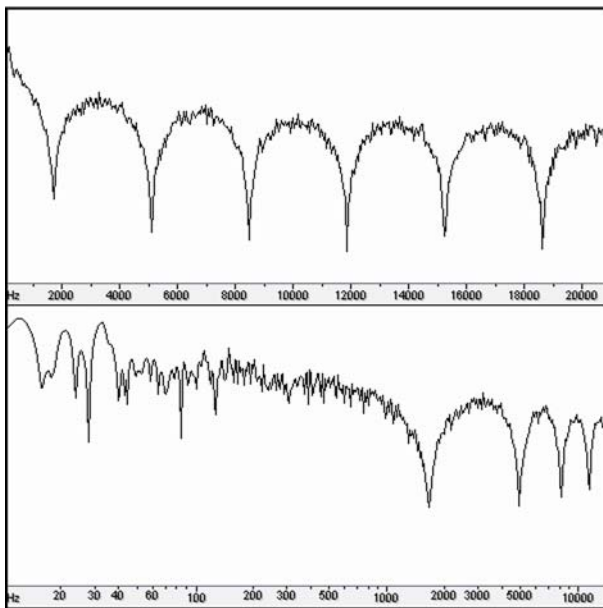
Hver enkelt refleksjons fysiske attributter kan beskrives med følgende tre kategorier: frekvensinnhold, amplitude og forsinkelse, hvor forsinkelsen defineres i forhold til ankomsten av den direkte lyden.

Skal man si noe om frekvensinnholdet til refleksjoner i det de ankommer lytteren må man starte med frekvensinnholdet til lydkilden selv. Det kan synes unødvendig å nevne noe så opplagt, men det er viktig i flere henseender. Først og fremst er det viktig fordi ulike lydkilder har ulik frekvenskarakteristikk, og dermed vil refleksjoner fra ulike lydkilder også ha ulik frekvenskarakteristikk. Med mindre det er helt statiske lydkilder vil den enkelte lydkilde også endre frekvenskarakteristikk over tid. Det viktige å merke seg i forhold til tidlige refleksjoner er at en lydkilde som ikke produserer lyd i et visst frekvensområde, heller ikke vil produsere refleksjoner i det samme frekvensområdet. I det følgende vil jeg trekke fram enkelte frekvensspesifikke trekk ved refleksjoner, og det trenger altså ikke være like betydningsfullt for alle lydkilder. I musikalsk sammenheng gjelder dette særlig lavere frekvensområder. Frekvensinnholdet til den enkelte refleksjon bestemmes videre av luftens absorpsjon, samt absorpsjonskoeffisienten til de overflatene den treffer og reflekteres fra på sin vei fram til lytteren. Absorpsjonskoeffisienten til ulike materialer er frekvensavhengig, noe som medfører at ulike frekvensområder vil bli absorbert i større eller mindre grad (Davis, Davis, 1997, s. 158). En refleksjon vil altså ikke være en eksakt kopi av den direkte lyden, men en filtrert versjon med frekvenskarakteristikk som påvirkes av materialet og utformingen av rommets overflater.

Det er også verdt å merke seg at de færreste lydkilder har en flat spredningskarakteristikk. Det innebærer at frekvensinnholdet i ulike refleksjoner fra en lydkilde allerede ved utgangspunktet kan være forskjellige. Det opplagte eksempelet på en lydkilde med delvis retningsbestemt utstråling er menneskets egen stemme. Lyden fra stemmebåndene våre strømmer hovedsakelig ut av en liten åpning på framsiden av hodet. Man opplever en stor klangforskjell avhengig av om man står foran eller bak dette hodet. Høytalere er et annet og, i denne sammenheng, kanskje mer aktuelt eksempel på det samme. En generell regel er at lydkilder tenderer til å bli mer og mer retningsstyrte med økende frekvenshøyde, mens spredningen av lavere frekvenser gjerne er mer rundstrålende (Rumsey, 2001, s. 4).

### 2.2.3 Kamfiltre og stående bølger

Ved komplekse lydkilder, som vi har i musikken, kan hver enkelt refleksjons innvirkning på den direkte lyden beskrives med et kamfilter. Et kamfilter er et filter som vekselvis demper og slipper gjennom frekvenser i jevne intervaller oppover frekvensspekteret. Kamfilteret har fått navnet sitt etter de karakteristiske, kamlignende taggene filteret har når man betrakter det på en lineær frekvensskala. Figur 2-1 viser den karakteristiske formen på den lineære frekvensskalaen, mens det på den logaritmiske frekvensskalaen blir stadig tettere mellom taggene.



**Figur 2-1** Øverst: kamfiltereffekt slik den framstår på lineær frekvensskala – Nederst: kamfiltereffekt slik den framstår på logaritmisk frekvensskala

Avstanden mellom topper og bunner i kamfilteret bestemmes av refleksjonens forsinkelse i forhold til den direkte lyden (Everest, 2001, s. 377). Tabell 2-1 viser noen eksempler på refleksjoners forsinkede ankomst versus kamfilterets karakteristikk.



Forsinkelse	Første null (tann) i kamfilter	Båndbredde mellom hver tann i kamfilter
0,5 ms	1000 Hz	2000 Hz
1 ms	500 Hz	1000 Hz
3 ms	167 Hz	333 Hz
8 ms	62,5 Hz	125 Hz
16 ms	31,25 Hz	62,5 Hz
25 ms	20 Hz	40 Hz

**Tabell 2-1:** Avstand mellom topper og bunner i kamfilter som følge av ulik forsinkelse

Som tabellen viser, får man mindre og mindre båndbredde mellom hver tann i filteret etter hvert som forsinkelse mellom direkte lyd og refleksjon øker. En konsekvens av dette er, som vi skal se nærmere på nedenfor, at refleksjoner med kortere forsinkelse har en større tendens til å påvirke lydens klangfarge.

At refleksjoners innvirkning på lyden er sterkt frekvensavhengig er spesielt tydelig i forbindelse med akustikken i rom som har dimensjoner som er sammenlignbare med lydens bølgelengde. Lyden beveger seg med en hastighet på omtrent 340 meter i sekundet, avhengig av temperatur, luftfuktighet og høyde over havet. Ved en lydhastighet på 340 meter i sekundet, vil lydbølgen til en sinustone på 100Hz være en hundredel av dette (den gjentas hundre ganger i løpet av et sekund), altså 3,4 meter. Bølgelengden til en sinustone på 50 Hz er sådan 6,8 meter, mens en sinustone på 440 Hz har en bølgelengde på 0,77 meter.

I et rektangulært rom hvor avstanden mellom to vegger er 3,4 meter vil det oppstå stående lydbølger mellom de to veggene ved frekvensene 50Hz, 100Hz, 150Hz, 200Hz, 250Hz, 300Hz osv. Stående lydbølger, eller romresonanser om man vil, har kort forklart den effekt at det generelt oppstår en forsterkning av den aktuelle frekvensen i rommet, mens den på enkelte punkter er kansellert. Frekvensen på 50Hz vil forsterkes maksimalt inne ved de to veggene, mens nullpunktet ligger midt mellom veggene. Samme type resonanser vil oppstå mellom de to andre veggene, og mellom gulv og tak. Er avstanden mellom de to andre veggene også 3,4 meter, vil man få resonanser ved akkurat de samme frekvensene. Konsekvensen er en ytterligere forsterkning av akkurat de samme frekvensene. Er avstanden eksempelvis 5,1 meter vil resonansene derimot oppstå ved 33,3Hz, 66,7Hz, 99,9Hz osv, og man vil få en jevnere fordeling av resonansene. Størrelsesforholdet mellom rommets lengde, bredde og høyde spiller dermed en avgjørende rolle for de akustiske forholdene i rommet. Generelt ønsker man en jevn spredning i de ulike resonansene som oppstår.

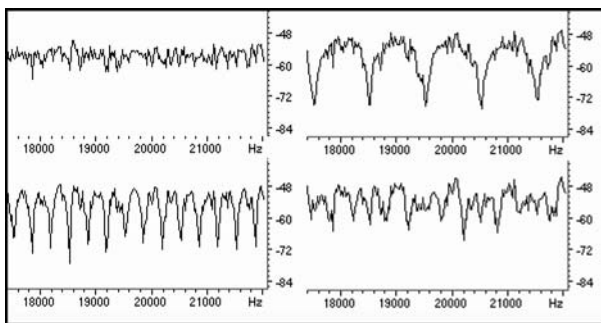
Resonansene som følge av at lyden beveger seg mellom to parallelle overflater kalles aksiale (aksial = som gjelder en akse). Disse er kraftigst og har derfor en tendens til å bli

vektlagt i litteraturen. Det vil også oppstå tangentielle og oblique resonanser. Tangentielle resonanser (tangential = som danner en tangent, altså som berører hverandre) innebærer en bevegelse mellom fire av rommets overflater og vil ha halvparten av den energien aksiale resonanser har. Oblique resonanser (oblikk = skjev) innebærer en bevegelse via alle seks overflatene i rommet og vil ha en fjerdedel av energien til aksiale resonanser. Tangentielle og oblique resonanser er dermed ikke fullt så kraftige, men har allikevel stor innvirkning på rommets frekvenskarakteristikk (Everest, 2001, s. 339).

I praksis innebærer dette at de fleste konsumentlytterrom og lydstudioer fungerer som resonante hulrom under cirka 300 Hz (Everest, 2001, s. 401). Stående bølger, med sine topper og nullpunkt, er spesielt en utfordring i rom med parallelle overflater, og parallelle overflater er et typisk trekk ved de fleste konsumentrom.

#### 2.2.4 Flere refleksjoner

I hvilken grad den enkelte refleksjon vil farge lyden, avhenger som nevnt av refleksjonens frekvensinnhold, amplitude og forsinkelse. Men refleksjonens posisjon og karakteristik i forhold til andre refleksjoner er også meget avgjørende. Negative effekter som følge av en enkelt refleksjon kan motvirkes ved tilføring av andre refleksjoner. Den innbyrdes avstanden mellom ulike refleksjoner har stor betydning. For å unngå uønsket farging av lyden vil man gjerne unngå at refleksjonene ankommer for tett eller for rytmisk (Halmrast, 1999, s. 6). En grad av tilfeldighet er normalt å foretrekke.



**Figur 2-2:** Øverst til venstre: Hvit støy – Øverst til høyre: Hvit støy mikset med 1 ms forsinket versjon av signalet – Nederst til venstre: Hvit støy mikset 3 ms forsinket versjon av signalet – Nederst til høyre: Hvit støy mikset med 1 og 3 ms forsinkede versjoner av signalet

Figur 2-2 viser endringene i frekvenskarakteristikk til en kompleks lyd (hvit støy) mikset med forsinkede versjoner av lyden. Kamfiltereffekten er meget tydelig ved de to enkeltrefleksjonene. Den er fremdeles synlig ved to refleksjoner, men den utflatende effekten av de to refleksjonene sammen er allerede tydelig.

Det er et par ting som er viktig å merke seg ved figur 2-2. For det første er refleksjonene eksakte kopier den originale lyden, både i frekvensinnhold og amplitude. I den virkelige verden vil kamfiltereffekten av enkle refleksjoner normalt være mindre grafisk tydelige enn i eksempelet. Videre illustrerer grafene en lydkilde med en viss kontinuitet og statisk preg, hvor det fortsatt finnes direkte lyd når refleksjonene ankommer. Det vil ikke oppstå noen kamfiltereffekt dersom man har en (meget) perkussiv lydkilde hvor den direkte lyden har opphørt før den første refleksjonen ankommer. Det kan allikevel oppstå kamfiltereffekter mellom ulike refleksjoner.

### **2.2.5 Antall tidlige refleksjoner**

Det er vanskelig å si noe generelt om hvor mange tidlige refleksjoner som oppstår i ulike rom, til det er verden alt for mangfoldig. De avgjørende faktorene vil være rommets størrelse og utforming, graden av diffusjon, lydkildens plassering og spredningskarakteristikk, og til slutt lytterens posisjon.

## **2.3 Psykoakustikken**

Jeg skal nå bevege meg over mot psykoakustiske effekter knyttet til tidlige refleksjoner. Det første som må sies i den forbindelse er at høresansen vår har en ganske annen innfallsvinkel i sin disseksjon av lyden, sammenlignet med vår akustiske tilnærming med oppdeling i direkte lyd og refleksjoner. Det innebærer ikke at et akustisk utgangspunkt i tilnærmingen til psykoakustikk er feilslått, men man skal være bevisst på grunnleggende forskjeller mellom de to. Mens akustikken kan betraktes som relativt statisk (har man først konstatert noe, stemmer det som regel ved gjentakelse), må psykoakustikken til tider tilskrives en dynamisk karakter.

Det opplagte eksempelet på psykoakustikkens dynamiske karakter er at alle mennesker er forskjellige. Høresansens evner baseres på økologisk erfaring. Man må anta at med det mangfoldet vi har i den klanglige verden vil også høresansens arbeid på vei opp til kognitive prosesser være preget av dette mangfoldet.

Et annet eksempel, som kanskje er mer viktig i denne sammenheng, er vår tilbøyelighet til å endre oss over tid. Mennesket innehar på enkelte områder en veldig kort klanglig hukommelse. Hørselen tenderer til å tilpasse seg de klanglige omgivelsene vi befinner oss i, men resetter seg relativt raskt ved opphold i lyden (Shinn-Cunningham, 2003,

s. 7). Det innebærer for eksempel at vår opplevelse av lyden på en konsert ikke nødvendigvis er lik ved starten og slutten av en låt.

Tidlige refleksjoner kan ha stor innvirkning på en lydkildes framtoning og karakteristikk. I tillegg bærer tidlige refleksjoner med seg viktig informasjon om en rekke romlige aspekter ved lydkilden, samt rommet både den og lytteren befinner seg i. Videre snakker man gjerne om ønskede og uønskede effekter knyttet til tidlige refleksjoner. Refleksjoner som er ønskelige på ett område, trenger ikke nødvendigvis å være det på andre.

Det er et langt lerret å bleke hvis en kartlegge alle perseptuelle effekter knyttet til tidlige refleksjoner. Jeg skal derfor hovedsakelig konsentrere meg om aspekter relevante for populærmusikkproduksjoner. Psykoakustisk forskning synes å ha en tendens til å fokusere på perseptuelle effekter av enkeltrefleksjoner, framfor mer helhetlige perspektiver, med den konsekvens at overførbarheten fra laboratorium til musikalske situasjoner ikke trenger være like stor.

### 2.3.1 Lydkilder

I de neste avsnittene skal jeg se på områder der tidlige refleksjoner kan spille en viktig rolle i vår opplevelse av ulike lydkilder.

#### *Lydstyrke*

Tidlige refleksjoner som ankommer lytteren opp til omtrent 50 ms etter den direkte lyden vil normalt oppleves som en del av den direkte lyden. Den totale lydenergien til stede i denne tidsperioden vil normalt bidra til en opplevd økning i lydstyrke (Everest, 2001, s. 74). De tidlige refleksjonenes innvirkning på opplevd lydstyrke er spesielt merkbar hvis man befinner ”off axis” (*norsk: bortenfor midtaksen*) i forhold til lydkilden. For eksempel hvis lytteren befinner seg til siden for en person som snakker. Den forsterkende effekten er tydeligst ved refleksjoner med relativt kort forsinkelse i forhold til den direkte lyden. En musikalsk konsekvens av de tidlige refleksjonenes varierende forsterkende effekt, er at man i orkestre gjerne plasserer instrumenter som kontrabasser og pauker relativt nær veggen. I psykoakustiske forsøk har Haas påvist økt lydstyrke i tidsperioden 0-30 ms (Bech, 1998, s. 437).

#### *Retning*

I hørselens analyse av retningen til en lydkilde gjelder primært ”regelen om første bølgefront” (*egen oversettelse av ”the law of the first wavefront”*), men flere har påvist at kraftige enkeltrefleksjoner kan påvirke retningsbestemmelsen. Blant annet har Olive og Toole påvist retningsendringer for enkeltrefleksjoner <10 ms (Bech, 1998, s. 437), mens Shinn-

Cunningham har påvist at retningsbestemmelsen fungerer langt bedre i reflekterende omgivelser sammenlignet med ekkofrie (Shinn-Cunningham, 2003, s. 3). Søren Bech (1998) har påvist at det kun er gulvrefleksjoner av første orden som alene vil påvirke spatiale aspekter ved reproduksjon i gjennomsnittlige konsumentrom.

I populærmusikkproduksjon benytter man seg gjerne av ulike ”precedence-effekter” (*norsk: forrang*) for å påvirke spatiale trekk i et lydbilde. Ordet precedence er relativt synonymt med regelen om første bølgefront. Ulike precedence-effekter innebærer i hovedsak tilføring av forsinkede versjoner av et lydsignal, og kan blant annet medføre en forskyving av den opplevde retningen til en lydkilde. Den velkjente Haas-effekten er en precedence-effekt. Haas-effekten vil bli beskrevet litt lenger nede.

#### *Avstand*

Tidlige refleksjoner er avgjørende i vår bedømming av avstanden til en lydkilde. Laterale refleksjoner (de som kommer inn fra siden) formidler avstand bedre enn frontale. I en brukermanual fra Lexicon (basert på David Griesingers arbeid) defineres det noen frekvensavhengige verdier for refleksjonenes innfallsvinkel for best formidling av avstand.

The optimal direction is frequency dependent. For frequencies below 700Hz the optimal direction is from the side (90 degrees from the front). At about 1500Hz the optimal direction is +/- 30 degrees from the front (the standard front speaker angle.) (Lexicon, 2001, s. 5-4)

#### *Forståelighet og klarhet*

Forståelighet (*engelsk: intelligibility*) er også et tema forbundet med tidlige refleksjoner. Emnet er kanskje tettest knyttet til forståelsen av tale, men er også aktuelt i musikalsk sammenheng. Problemer rundt forståelighet knyttes helst til senere refleksjoner og styrkeforholdet mellom direkte lyd og reflektert lyd. Refleksjoner i tidsregionen 50-150 ms er spesielt problematiske, men også kamfiltereffekter som følge av tidlige refleksjoner kan trekke forståeligheten i negativ retning.

Forståeligheten er beslektet med det akustiske begrepet ”clarity” (*norsk: tydelighet/klarhet*), som er et mål på den samlede lydenergien før og etter et gitt tidspunkt. Høyere verdier før det gitte tidspunktet innebærer økt grad av tydelighet. I forbindelse med forståelsen av tale opererer man gjerne med Clarity50, hvor man da ønsker at hovedmengden av lydenergi fra en lydkilde skal ankomme lytteren innen 50 ms etter den direkte lyden. I musikken har man ikke fullt så strenge krav til tydelighet, og man opererer derfor gjerne med Clarity80. Clarity-begrepet sier kun noe om lydenergi før og etter et gitt tidspunkt, ikke hvordan lydenergien er fordelt innenfor de to tidsperiodene. Det sier heller ingenting om frekvensinnholdet i lyden.

### *Utstrekning*

Refleksjoner opp til 50-80 ms etter den direkte lyden har vanligvis den effekt at de utvider vår opplevelse av en lydkildes utstrekning i både bredde og dybde (Rumsey, 2001, s. 33). Apparent Source Width (ASW) er et akustisk begrep knyttet til opplevelsen av en lydkildes bredde. ASW synes å ha liten relevans i akustikken til små rom. Rumsey påpeker at det i reprodusert lyd kan være vanskelig å avgjøre om en lydkilde er bred eller om den bare diffus og vanskelig å plassere (Rumsey, 2001, s. 37).

Den mye brukte Haas-effekten, som ble nevnt ovenfor, endrer blant annet opplevelsen av en lydkildes bredde i reprodusert stereo. Haas påviste at en forsinket versjon av et signal, eller med andre ord en refleksjon, i tidperioden 5-35ms må forsterkes mer enn 10dB over den direkte lyden før den oppleves som et ekko (Everest, 2001, s. 74). Grenseverdiene i praktisk bruk er nok litt strengere og avhenger sterkt av lydkildens karakteristikker.

### *Klangfarge*

Tidlige refleksjoners påvirkning av lydens klangfarge er tett forbundet med de kamfiltereffekter som oppstår mellom den direkte lyden og summen av refleksjonene som treffer lytteren. Klangfargeendringer kan være av både ønsket og uønsket art. Refleksjoner under 10 ms er generelt assosiert med uønskede effekter. Refleksjoner som ankommer under 10 ms etter den direkte lyden oppstår gjerne i mindre rom, og negative kamfiltereffekter er derfor et karakteristisk trekk i mange kontrollrom og konsumentrom.

I et prosjekt for å forbedre de akustiske forholdene i Oslo Konserthus, påviste Tor Halmrast at lyden fra et orkester på en scene får en boksaktig klangfarge (*tysk: Box-Klangfarbe*) når en distinkt refleksjon i tidsrommet 5-20 ms skaper et klart kamfilter med båndbredde mellom hver tann som sammenfaller med ørets kritiske bånd (Halmrast, 1999, s. 4). Halmrast påviste videre at man ved tilføring av flere refleksjoner i en tilfeldig orden vil flate ut klangfargeendringene (Halmrast, 1999, s. 4).

Antallet lydkilder er i så måte en avgjørende faktor i tidlige refleksjoners farging av lyden. Et symfoniorkester består av en rekke instrumenter. Hvert instrument i orkesteret har en unik posisjon i rommet, med den konsekvens at hvert instrument vil generere refleksjoner som vil avvike fra de andre instrumentenes refleksjoner, både i tid og frekvenskarakteristikk. Alle de små nyansene i refleksjoner vil ha en utflatende effekt på klangfargeendringer. Dette er et spesielt viktig punkt i forbindelse ved reproduksjon i stereo. Ved en reproduksjon i stereo vil lyden fra alle instrumenter strømme ut fra to høyttalere, og man vil dermed "miste" en god del utflatende refleksjoner. Jeg snakker her om den lyden som oppstår i rommet hvor

man lytter til musikken reproduert i høyttalere, for eksempel hjemme hos konsumenten. Det er selvsagt en forenkling av hva som skjer ved reproduksjon av en innspilling av et orkester. Det viktige å merke seg er at jo flere lydkilder man har, desto mindre utsatt vil lyden være for klangfargeendringer. Dette innebærer også at reproduksjon i mono vil være mer utsatt for klangfargeendringer enn reproduksjon i stereo.

#### *Ulike lydkilder*

De tidlige refleksjonenes innvirkning på vår opplevelse av lydkilder avhenger naturlig nok av karakteristikken til den enkelte lydkilde. Generelt kan man si at refleksjoner fra transientrike lydkilder i større grad vil påvirke opplevelsen vår, sammenlignet med lydkilder som inneholder færre transienter. En rekke faktorer i det samlede preget til en lydkilde, i kombinasjon med rommet der lydkilden genererer lyd, er det som til slutt avgjør hvorvidt refleksjonene påvirker opplevelsen. Vitenskapen har, som vi senere skal se, pekt på en rekke grenseverdier for når hørselen opplever refleksjoner. Det er i bunn og grunn karakteristikken til lydkilden som avgjør om refleksjonene vil overstige disse grenseverdiene eller ikke.

Griesinger stadfester blant annet at det kun er fluktueringer i tid og amplitude mellom ørene i løpet av startfasen til lydhendelser som påvirker lydkildens bredde. Fluktueringer i tid og amplitude forkortes henholdsvis ITD = Interaural Time Difference og IID = Interaural Intensity Difference (Griesinger, 2005a, s. 2) Fluktuering i IID og ITD oppstår blant annet når refleksjoner ankommer lytteren. Det skjer også når lydkilde eller lytter flytter seg. Det innebærer at tidlige refleksjoners innvirkning på opplevd bredde avhenger av brattheten i starten av en lydkildes omhyllingskurve. Sagt med andre ord, håndklapp påvirkes generelt mindre enn fiolin strøket med bue. Selv om man opererer med refleksjoners innvirkning på ASW helt opp i 80 ms, er det for vokal og en rekke andre soloinstrumenter gjerne kun refleksjoner før 10 ms som er aktuelle. Refleksjoners innvirkning på både retningsbestemmelse, samt emner som omhylling og romlighet, påvirkes på samme måte av lydkildenes karakteristik.

### **2.3.2 Romopplevelser**

Opplevelsen av romlige aspekter assosieres gjerne med de senere refleksjonene, men tidlige refleksjoner spiller en rolle på flere områder. I løpet av livet bygger hørselen opp en enorm erfaring fra ulike rom. Hørselens tolking av omgivelsene baserer seg på denne erfaringen. Hørsansen er vanligvis assistert av synssansen, så de færreste av oss reflekterer over den store kapasiteten hørselen har til å identifisere omgivelsene vi befinner oss i.

Tidsforsinkelsen mellom direkte lyd og tidlige refleksjoner er relativt proporsjonal med størrelsen på rommet, og sammen med refleksjonenes karakteristikk og amplitude gir dette hørselen viktig informasjon om rommets størrelse og utforming.

Shinn-Cunningham har påvist at vi i hverdagslige omgivelser har problemer med å identifisere vår egen posisjon i rommet, men at vi ved posisjoner tett inntil vegger, hvor det ene øret er rettet inn mot veggen og lydkilden er relativt langt borte, har litt lettere for å bedømme posisjonen (Shinn-Cunningham, 2003, s. 5). En posisjon nært en vegg vil normalt gi en kraftig tidlig refleksjon med relativt kort forsinkelse.

### *Romlighet*

Opplevelsen av romlighet (*engelsk: spaciousness*) og den beslektede opplevelsen av omhylling er viktige begreper i både konsertsaler og reprodusert lyd. Enkelte setter likhetstegn mellom opplevelsen av romlighet og opplevelsen av omhylling, men dette er problematisk i omtalen av lydbildet som oppstår ved i reproduksjon i stereo. Jeg anser derfor de to opplevelsene som ulike.

Romlighet er en opplevelse av rommet lydkilden/lytteren befinner seg i. Romlighet beskrives vanligvis ut fra rommet lytteren befinner seg i, gjerne med instrumenter som produserer lyd i det samme rommet. I reprodusert musikk derimot, vil opplevelsen av romlighet knyttes til lydbildets formidling av rommet der instrumentene befinner seg. Opplevelsen av romlighet er mest relatert til sene refleksjoner etter 150ms, men laterale tidlige refleksjoner er også viktige. Olive og Toole har påvist effekten for refleksjoner ved 10-40ms (Bech, 1998, s. 437).

Omhylling knyttes til lytterens opplevelse av å være en del av/inkluderes i lydbildet. De samme grunnleggende kriteriene, laterale refleksjoner og refleksjoner etter 150ms, gjelder for omhylling som for romlighet. Frekvenser under 500Hz er spesielt viktige for opplevelsen av omhylling. Samtidig er den optimale innfallvinkelen på laterale refleksjoner frekvensavhengig. Ved frekvenser under 700 Hz er den optimale innfallvinkelen 90°. Frekvensavhengigheten innebærer at det i et vanlig høyttaleroppsett  $\pm 30^\circ$  kun er frekvensinnhold over 1000Hz som skaper omhylling (Griesinger, 1998b, s. 5)

### *Dybde*

Opplevelsen av dybde i det lydbildet som oppstår i reprodusert stereo kan være knyttet til avstanden til lydkilden lengst borte, men er ikke bundet til den. Griesinger anbefaler bruken av tidlige refleksjoner opp til 50ms for å simulere dybde (Rumsey, 2001, s. 33).

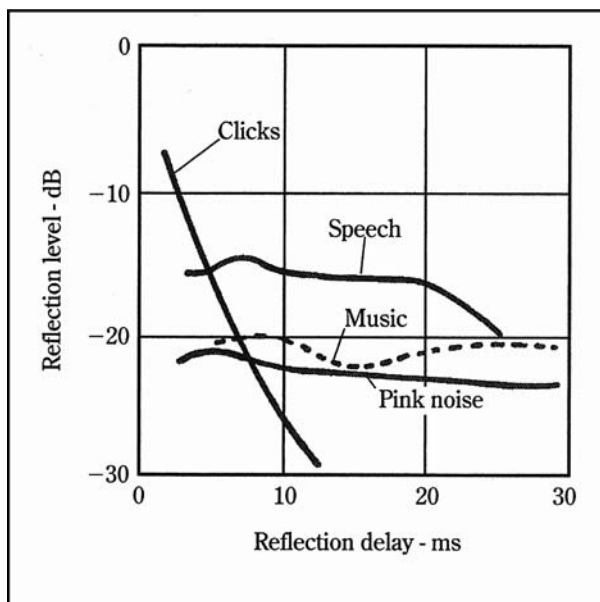


### 2.3.3 Grenseverdier

Den grunnleggende forutsetningen for at tidlige refleksjoner skal påvirke vår persepsjon av lyden er at de enten enkeltvis eller samlet inneholder tilstrekkelig energi. Psykoakustiske forsøk har avdekket en rekke grenseverdier for når refleksjonene er uhørbare, når de påvirker opplevelsen av den direkte lyden og når de oppleves som egne lydhendelser (ekko). Grenseverdier for klangfargeendringer er generelt lavere enn grenseverdier for spatiale trekk (Bech, 1998, s. 439). Maskering er et nøkkelord. Maskering av refleksjoner som følge av den direkte lyden er avgjørende for refleksjonenes hørbarhet. I musikalsk sammenheng er maskering fra andre lydkilder også meget viktig. Figur 2-3 er sakset fra *The Masters Handbook of Acoustics* og viser grenseverdier for persepsjonen av ulike lydkilder.

The so-called "early reflections" are only audible after a note starts. They are sometimes audible as a change in localization or timbre while a note is held, but in general they affect perception most strongly only after a note ends (Lexicon, 2001, s. 5-3).

Start- og stopptidspunktet mellom lydhendelser er nært knyttet til maskeringseffekter og er et grunnleggende kriterium i vår persepsjon av romklang generelt. Sitatet ovenfor kan for øvrig knyttes direkte til det Griesinger har kalt tidlig spatiale inntrykk (ESI).



**Figur 2-3:** "Absolute thresholds of perception of reflections of different types of signals, ranging from 2/second clicks (noncontinuous), to pink noise (continuous). The closeness of pink noise to Mozart music gives assurance that pink noise is a reasonable surrogate for music measurements. After Olive and Toole. *Journal of the Audio Engineering Society*" Figur og tilhørende tekst er kopiert fra *The Master Handbook of Acoustics* (Everest, 2001, s. 358)

Kamfiltre har vært et sentralt tema dette kapitlet. Grenseverdiene for ulike kamfiltereffekter avhenger i stor grad av samsvar mellom kamfilterets båndbredde mellom hver tann og ørets kritiske bånd.

Graden av diffusjon spiller også en grunnleggende rolle i vår persepsjon av romklengen og kan ha stor innvirkning på ulike grenseverdier, blant annet opplevelsen av distinkte ekko. Et godt eksempel finner man i impulsresponsen til den anerkjente konsertsalen Musikverein i Wien. Etter cirka 70 ms ankommer det en refleksjon som er så kraftig at den ville oppleves som et distinkt ekko om ikke refleksjonene før var så fyldige og bra distribuert i tid (Halmrast, 1999, s. 4).

Til slutt må det nevnes at refleksjonenes frekvenskarakteristikk og fase i forhold til den direkte lyden, er viktig med hensyn til både forsterkning, klangfarge og endringer av spatiale trekk. Bech har påvist at den spektrale energien over 2 kHz avgjør refleksjonens viktighet for spatiale aspekter i reprodusert lyd (Bech, 1998, s. 444).

## 3 Musikkproduksjon

I forrige kapittelet beskrev jeg grunnleggende elementer knyttet til tidlige refleksjoner. Jeg skal nå forsøke å kople denne kunnskapen til populærmusikkproduksjon. Jeg vil starte med å se på praktiske konsekvenser i tilknytning til de tidlige refleksjoner som fanges opp av mikrofoner i innspillingsrommet, og deretter utfordringer knyttet til reproduksjonen i lydstudioets kontrollrom. I fortsettelsen der jeg ser på mikseprosessen kommer den kunstige romklangen inn i bildet. Etter et grundig blick på kunstig romklang, vil jeg avslutte kapittelet med noen eksempler på lydteknikere/produsenters praktiske bruk av kunstig skapt romklang.

### 3.1 Innspillingen

I hvilken grad tidlige refleksjoner vil prege lyden som fanges opp av mikrofoner i et lydstudio, avgjøres hovedsakelig av tre faktorer:

1. Innspillingsrommets utforming
2. Lydkildens karakteristikk og posisjon i rommet
3. Mikrofonenes karakteristikk og posisjon i rommet

Av disse tre kan innspillingsrommets utforming betraktes som en premissleverandør for de to andre, da grunnpremissene for refleksjonenes karakteristikk er å finne i utformingen av rommet og graden av absorpsjon i rommets overflater. Befinner man seg for eksempel i et ekkofritt rom vil refleksjoner fra rommet være nærmest ikke-eksisterende, og verken lydkildens eller mikrofonens posisjon og karakterstikk vil spille nevneverdig rolle med hensyn til tidlige refleksjoner. På samme måte vil et lite, høyt reflekterende rom ubønnhørlig prege lyden som fanges opp.

I dagens flerspors populærmusikkopptak ønsker man generelt at refleksjoner fra innspillingsrommet skal farge den innspilte lyden i minst mulig grad. Dette gjelder spesielt med tanke på uønskede klangfargeendringer, men også refleksjoner som formidler spatiale trekk mellom lydkilde og innspillingsrom. Det er selvsagt ingen allmenngyldig regel, det finnes utallige eksempler på det motsatte, men kan allikevel anses som den generelle trenden som preger majoriteten av populærmusikkinnspillinger. En konsekvens av dette er at man i profesjonelle lydstudioer favoriserer store innspillingsrom hvor man har bedre kontroll med de akustiske forholdene, særlig i tilknytting til romresonanser og tidlige refleksjoner. Gilford slo fast i 1959 at studioer med størrelse under cirka 1500 kubikkfot er så utsatt for lydfarging at de er upraktiske (Everest, 2001, s. 401). 1500 kubikkfot (42,5 m<sup>3</sup>) tilsvarer for eksempel et rom på 3 x 3,5 x 4 meter, eller et rom på 18 m<sup>2</sup> og en takhøyde på 2,4 m.

Det mest påfallende med denne grenseverdien er antageligvis at man her muligens berører et av de tydeligste skillene mellom hjemmestudioer og mer profesjonelle etter den digitale revolusjonen. Profesjonelle studioer er generelt i besittelse av større rom, men det finnes også en hel del mindre rom, selv i profesjonelle studioer. Mindre rom benyttes gjerne til opptak av vokal i situasjoner der vokalopptakene blir gjort i samspill med andre instrumenter. I profesjonell sammenheng er det også interessant med tanke på den trenden man har sett i norsk musikk i noen år, hvor fremtredende artister flytter ut av det profesjonelle lydstudioet og inn i mer tiltalende omgivelser for å gjøre opptakene til en produksjon der. Et godt eksempel er det norske bandet Vamp, som har gjort hovedmengden av opptakene sine etter år 2000 på Ryvarden fyr utenfor Haugesund. Vamps produksjoner kan i så måte betraktes som profesjonelle hjemmestudioinnspillinger, hvor man etter opptakene beveger seg inn i det profesjonelle lydstudioet for nedmiksing og etterproduksjon. Det er ikke få profesjonelle produksjoner i dag som har fulgt en slik prosess.

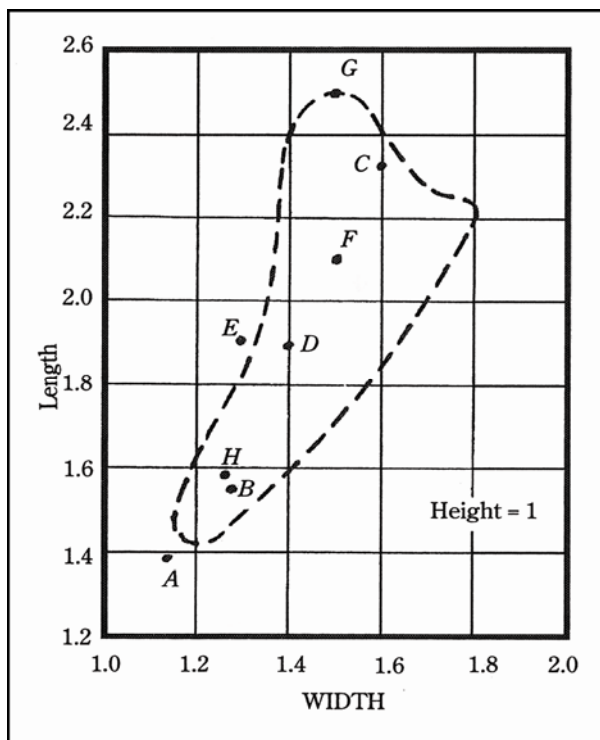
Gyldigheten av Gilfords grenseverdi kan absolutt diskuteres, både med hensyn til mikrofonplassering, ulike typer lydkilder og ikke minst innspillingsrommets utforming. Det må nevnes her at Gilford kun har klassifisert romstørrelsen som upraktisk, ikke ubrukelig. Jeg skal først se litt mer på innspillingsrommets utforming, og kommer tilbake til de to andre lenger nede.

Vi har sett at uønsket lydfarging gjerne er en konsekvens av stående lydbølger og kamfiltereffekter fra tidlige refleksjoner. Stående lydbølger er først og fremst en utfordring i frekvensområdet under 300 Hz. I hvilken grad resonansene vil være kraftige nok til å påvirke vår opplevelse av lyden, avhenger blant annet av rommets absorpsjon av de ulike resonansene. Et rom med høy absorpsjon i lavere frekvensområder vil generelt være mindre

utsatt for uønsket farging av lyden. I lydstudioer oppnår man gjerne denne type absorpsjon ved å sette inn spesialutformede absorbenter for å kontrollere ulike frekvensområder. Absorbenter kan dempe frekvensområder og til en viss grad jevne ut frekvensresponsen til rommet, men det er allikevel vanskelig å få til en jevn frekvensrespons i mindre rom.

Den enkelte resonans' innvirkning på lyden avgjøres videre av resonansens båndbredde, samt avstanden til andre resonanser. Resonansenes båndbredde er omvendt proporsjonal med etterklangstiden i rommet (Everest, 2001, s. 337), og etterklangstiden i et rom avhenger som kjent av graden av absorpsjon i rommet. Jo kortere etterklangstid, desto større vil båndbredden til hver resonans være. Båndbredden i et typisk lydstudio er omtrent 5 Hz (Everest, 2001, s. 337).

Fordelingen av resonanser i lavere frekvensområder avhenger som nevnt av størrelsesforholdet mellom rommets lengde, bredde og høyde. Ulike forskere har angitt ulike tall for hva som er de gunstigste forholdstallene mellom dimensjonene i rektangulære rom. Flere oppgir helt distinkte forholdstall, som Boner (1,00:1,26:1,59) og Volkmann (1,00:1,5:2,5) (Everest, 2001, s. 277). Mer anvendelig er kanskje det såkalte Bolt-området som definerer et litt større område som akseptabelt (se figur 3-1).



**Figur 3-1:** A chart of favorable room dimensional ratios to achieve uniform distribution of modal frequencies of a room. The broken line encloses the so-called "Bolt-Area." Bokstavene "B" og "G" refererer til henholdsvis Boner og Volkman, (resten av bokstavene er utelatt fra oppgaven) Figur og engelsk tekst er kopiert fra *The Master Handbook of Acoustics* (Everest, 2001, s. 276)

Hvorvidt en tidlig refleksjon vil farge vår opplevelse av lyden i et opptak, påvirkes kanskje først og fremst av lydkilden og/eller lytterens nærhet til reflekterende overflater. Det er også den opplagte nærheten til reflekterende overflater i Gilfords rom under 1500 kubikkfot som gjør dem så utsatt for lydfarging. Ved en posisjon mot midten av et rom, vil nærmeste reflekterende overflate vanligvis være enten gulv eller tak. En praktisk konsekvens av dette er at en viss takhøyde er å foretrekke i innspillingsrom.

### **3.1.1 Lydkildens posisjon og karakteristikk**

Et meget viktig aspekt i denne diskusjonen er det faktum at frekvenser som ikke genereres av en lydkilde, heller ikke vil eksistere som refleksjoner. Det er nærliggende å trekke fram tverrfløyte som eksempel, men det er antakeligvis mer matnyttig å ta utgangspunkt i menneskestemmen. Gjennomsnittlig verdi for grunnfrekvensen (*engelsk: fundamental frequency*) til den mannlige stemme ligger på 120 Hz, mens den kvinnelige ligger på 210 Hz (Traunmüller, Eriksson 1995, s. 1). En mannlig bass vil produsere i grunnfrekvenser ned mot 75 Hz, men gjennomsnittet ligger altså en del høyere. Dette betyr at menneskets stemme sjelden vil gi liv til de laveste resonansene i rom, med unntak av veldig små rom. Det er blant de laveste frekvensene man finner størst avstand mellom de ulike resonansene, og man kan dermed si at menneskestemmen ikke er fullt så utsatt for uønsket lydfarging sammenlignet med instrumenter som produserer lavere frekvenser. Dette betyr ikke at stående lydbølger ikke er en utfordring ved innspilling av vokal i mindre rom, det illustrerer kun at det er ulikheter mellom lydkilder.

Den tydeligste, og gjerne minst ønskelige musikalske konsekvensen av stående lydbølger oppstår når enkelte tonehøyder blir mye kraftigere enn de andre. Det er et utpreget, ofte stort problem i konsertsammenheng, og også et kjent fenomen fra musisering i baderom.

Påvirkning av enkelte tonehøyder kan også være en konsekvens av kamfiltereffekter fra tidlige refleksjoner. For eksempel vil en refleksjon 2 ms etter den direkte lyden danne et kamfilter som henholdsvis demper og forsterker annenhver partialtone til tonehøyden C4 (261,63 Hz) (se tabell 3-1).

Bunn i filter	Partialtone oddetall	Topp i filter	Partialtone partall
250 Hz	262 Hz (1)	500 Hz	523 Hz (2)
750 Hz	785 Hz (3)	1000 Hz	1047 Hz (4)
1250 Hz	1308 Hz (5)	1500 Hz	1570 Hz (6)
1750 Hz	1831 Hz (7)		

**Tabell 3-1:** Sammenfall mellom kamfilter som følge av 2 ms forsinkelse og partialtoner i tonen C4. Refleksjonens innvirkning på klangfargen er åpenbar

En forsinkelse på 2 ms innebærer at refleksjonen vil ha reist 70 centimeter lenger enn den direkte lyden når den treffer lytteren, og det igjen betyr at lydkilde og/eller lytter befinner seg relativt nær en reflekterende overflate. Vi har sett at uønskede kamfiltereffekter er spesielt knyttet til refleksjoner som ankommer mindre enn 10 ms etter den direkte lyden. En forsinkelse på 10 ms innebærer at en refleksjon har reist 3,44 meter lenger enn den direkte lyden. I større rom er dette ingen uvanlig reiselengde selv for refleksjoner av første orden, men i mindre rom vil de første refleksjonene ha en kortere reise. Nok en gang ser vi altså at romstørrelse spiller en viktig rolle, men spørsmålet i denne delen er hvordan dette forholder seg til innspilling av populærmusikk. Hvis vi fortsetter med vår egen stemme som eksempel, er det flere faktorer man har kontroll over under en innspilling. Sangerens posisjon i rommet vil, som vi nå vet, være avgjørende for refleksjonenes karakteristikk. Sangerens avstand til veggene i rommet kan endres med enkelhet, og en posisjon nær midten av rommet kan synes å være gunstigst. Avstanden til gulv og tak derimot, er relativt konstant. Selv i mindre rom, vil gulv og tak gjerne være de to nærmeste overflatene. Avstanden til gulvet er for øvrig den faktoren som er noenlunde konstant uansett romstørrelse, og første refleksjon som når lytteren stammer gjerne fra gulvet. Videre har gulv og tak gjerne en glattere overflate enn veggene. Refleksjoner fra glatte overflater er generelt mindre utsatt for diffusjon (spredning) enn refleksjoner fra mer ujevne overflater. Diffusjon innebærer, kort forklart, en større spredning av lydenergien, og refleksjoner fra glatte overflater er derfor gjerne mer distinkte.

### 3.1.2 Mikrofonens posisjon og karakteristikk

Hvis vi nå skal nærme oss de faktiske konsekvensene av tidlige refleksjoner i innspillingsammenheng vil aspekter knyttet til mikrofonbruk begynne å gjøre seg gjeldende. Den absolutt viktigste faktoren i spørsmålet om tidlige refleksjoner vil ha betydning for vår opplevelse av lyden er refleksjonenes amplitude i forhold til den direkte lyden. Det grunnleggende nivåforholdet mellom direkte lyd og refleksjoner er å finne i reiselengden. Den desidert mest brukte mikrofonteknikken i populærmusikkinnspillinger er såkalt nærmikking, som enkelt og greit betyr at mikrofonen plasseres nært lydkilden. Den store fordelen med

nærmikking er nettopp at man øker forholdet i reiselengde mellom den direkte lyden og refleksjonene. For eksempel vil første refleksjon fra en sanger på 174,9 cm som står 30 centimeter fra mikrofonen og har gulvet som nærmeste overflate, ha en amplitude som er minst 21 dB svakere enn den direkte lyden. Er takhøyden i rommet 240 cm, vil første refleksjon være minst 14 dB svakere. Øker man avstanden til mikrofonen til 60 cm, vil nivåforskjellen minst være henholdsvis 9 og 15 dB. Disse tallene tar kun utgangspunkt i ulikheter i reiselengde. Den reelle nivåforskjellen vil også påvirkes av overflatenes absorpsjon, og vil dermed normalt være større.

Figur 2-3 viste grenseverdier for vår opplevelse av refleksjoner ved ulike typer lydsignaler. Man kan anta at grenseverdiene for opptak av en enkelt sangstemme ligger et sted mellom verdiene for tale og verdiene for musikk, men det er før man plasserer stemmen i en musikalsk helhet sammen med andre instrumenter. Lytter man til opptakene av stemmen alene vil første refleksjon kunne spille en rolle i alle fall i noen av eksemplene ovenfor. Det er derimot mer usikkert om man vil kunne oppleve det samme når sangopptaket inngår i en større musikalsk helhet. I en større musikalsk helhet vil refleksjonens nivå stå i relasjon til et helhetlig lydbilde, og gjerne være et offer for maskering som følge av lyden fra andre instrumenter. Det er umulig å si noe konkret i et så lite spesifikt eksempel som dette er, til det er det alt for mange faktorer som spiller inn, men en kan anta at effekter av refleksjonene i dette eksempelet vil tendere til å være umerkbare.

Ved opptak av vokal i populærmusikkproduksjoner plasserer man gjerne sangeren enda tettere på mikrofonen. Avstander ned mot noen få centimeter er ikke uvanlig, med den konsekvens at refleksjonenes relative lydstyrke vil være enda lavere.

Det må nevnes her at i praksis avgjøres nok avstanden mellom sanger og mikrofon i større grad av den klangfargeendring som oppstår som følge av nærhetseffekten (*engelsk: proximity effect*), enn refleksjonenes relative lydstyrke. Nærhetseffekten er en konsekvens av fysikken til retningsstyrte mikrofoner, og innebærer en forsterkning av lavere frekvenser. Den oppstår ved cirka 30 centimeters avstand til mikrofonen, og tiltar omvendt proporsjonalt med avstanden. Effekten gir gjerne en utpreget fylde til vokal, og er derfor ofte en klangfargeendring man tilstreber ved opptak av vokal til populærmusikk. En nær avstand til mikrofonen er dermed gjerne en vinn-vinn-situasjon ved opptak av vokal i populærmusikk. Man får en positiv klangfargeendring og tar samtidig opp et signal med en veldig nærhet, som er et godt utgangspunkt før eventuell manipulasjon i mikseprosessen. De akustiske utfordringene ved nærmikking av vokal er således sterkere knyttet til stående bølger, enn effekter av tidlige refleksjoner, forutsatt at gulvet er nærmeste overflate.



Vokal er en velegnet lydkilde til nærmikking fordi så godt som all lyd strømmer ut av en liten åpning, og det er i tillegg relativt enkelt å plassere denne åpningen i god avstand til reflekterende overflater. Man oppnår dermed et bra resultat ved å plassere en enkelt mikrofon tett inntil åpningen. Bildet endrer seg når man beveger seg over på andre lydkilder.

En del instrumenter har en fysisk utforming som gjerne tvinger dem nærmere gulvet. De fleste instrumenter har i tillegg en mer kompleks spredningskarakteristikk som gjerne framtvinger en større avstand og/eller flere mikrofoner for å oppnå tilfredsstillende lyd kvalitet. En innføring av flere mikrofoner for å ta opp et instrument gjør bildet mer komplisert. Man kan naturlig nok forutse refleksjonenes innvirkning på signalet til den enkelte mikrofon, på samme måte som ved vokaleksempelen ovenfor. Den store utfordringen ligger i kombinasjonen av signalene fra flere mikrofoner, og interaksjonen mellom disse i reproduksjon.

I signalet fra en enkelt mikrofon vil man ikke kunne skille mellom frontale og laterale refleksjoner. Man kan i persepsjonssammenheng betrakte alle refleksjoner som frontale, med de begrensninger det har for vår opplevelse av spatiale trekk. Ved innføringen av en mikrofon til åpner man for muligheten til å reprodusere laterale refleksjoner i større grad, forutsatt at reproduksjonen foregår i stereo og signalene ikke panoreres likt. Dette er et stort tema jeg ikke skal gå i dybden på, men det må nevnes at mikrofonenes posisjon i forhold til hverandre har stor betydning for lyden. Dette gjelder spesielt i situasjoner hvor signalet fra to eller flere mikrofoner sendes til samme høyttaler i reproduksjon, men gjelder også ved full panorering. Det vil alltid være et visst avvik i ankomsttiden til den direkte lyden mellom hver av mikrofonene. I bunn og grunn er det hovedpoenget med bruk av flere mikrofoner, men det viktige her er at man kan betrakte avviket på samme måte som refleksjoner. Avhengig av lengden på forsinkelsen vil et forsinket signal fra en mikrofon kunne medføre kamfiltereffekter på lik linje med refleksjoner.

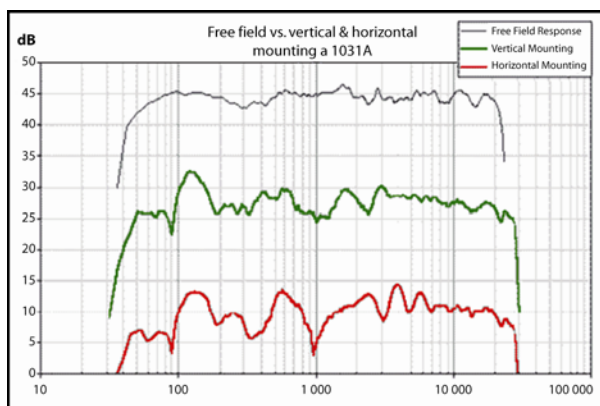
Så langt kan det virke som man i populærmusikkproduksjon utelukkende forsøker å minimere refleksjoner fra rommet i størst mulig grad. Det er ikke nødvendigvis slik. I mange sammenhenger ønsker man å fange opp mer av responsen fra innspillingsrommet. Ved bruk av flere mikrofoner har man som nevnt mulighet til å reprodusere refleksjonenes innfallsvinkel i større grad. Man har dermed en større mulighet til å formidle opptaksrommets karakteristikk. Det finnes en rekke godt utprøvde mikrofonteknikker som er velegnet til å gi en realistisk reproduksjon av rommet i stereo. For å oppnå en realistisk reproduksjon av rommet forutsettes det en viss grad av nøyaktighet med hensyn til mikrofonenes plassering.

Til slutt må det sies at i populærmusikkproduksjon er det ikke alltid slik at idealet er en realistisk reproduksjon, selv når man ønsker å formidle romfølelse fra innspillingsrommet. I så måte har man muligheten til ved mikrofonplassering alene å formidle både tilnærmet realistiske og helt urealistiske rom.

### 3.1.3 Kontrollrom

De akustiske forholdene i kontrollrommet er avgjørende for mange aspekter ved en populærmusikkproduksjon. Det er i kontrollrommet lydtekniker og produsent gjør sine valg, og de kan i uheldige omstendigheter basere sine valg på feilaktig klanglig informasjon. I profesjonelle lydstudioer legges det derfor normalt stor innsats i å optimalisere lytteforholdene. Selv mange profesjonelle studioer har et begrenset areal til rådighet, noe som innebærer at kontrollrom ofte faller inn under kategorien mindre rom. Og som vi allerede har sett, er det en del akustiske utfordringer knyttet til rom med liten størrelse.

Det er i tillegg noen relativt faste faktorer i kontrollrom som gjør dem spesielt utfordrende med tanke på uønskede effekter knyttet til tidlige refleksjoner. Den mest prominente ligger i interaksjonen mellom miksepulten og såkalte nærfeltmonitører som vanligvis plasseres på toppen av, eller like bak miksepulten. Miksepulten er i praksis en stor, høyt reflekterende overflate og lydteknikeren sitter vanligvis tett inntil den. En tidlig refleksjon fra miksepulten vil innebære en kamfiltereffekt. Man klarer å redusere kamfiltereffekten noe ved å konstruere miksepulter som heller, og ytterligere ved å plassere monitorene i best mulig posisjon, men en viss kamfiltereffekt synes unngåelig. Figur 3-2 viser frekvensresponsen til en nærfelthøytaler i henholdsvis ekkofritt rom, stående på stativ like bak miksepulten og liggende oppå miksepulten. Målingen er utført av høytalerprodusenten Genelec.



**Figur 3-2:** Grafene viser frekvensresponsen til en nærfeltmonitor i henholdsvis ekkofritt rom (øverst), stående på stativ like bak miksepulten og liggende oppå miksepulten. Figuren her kopiert fra nettsidene til Genelec (Genelec, 2006).

Et viktig aspekt i konstruksjonen av kontrollrom er å unngå at tidlige refleksjoner i kontrollrommet maskerer aspekter ved refleksjonene som oppstår i innspillingsrommet. Beranek definerte i sin tid begrepet "the initial time-delay gap" som er et mål på forsinkelsen mellom ankomsten av den direkte lyden og ankomsten av den første tidlige refleksjonen, og påviste sammenhengen mellom forsinkelsen og opplevd lyd kvalitet konsertsaler (Everest, 2001, s. 429). Beraneks funn var at konsertsaler med god lyd hadde en tydelig forsinkelse på rundt 20 ms. I kunstige romklangsimulatorer benytter man for øvrig begrepet "predelay" om den samme forsinkelsen. Davis relaterte dette til lydstudioer og påviste at variasjoner i denne forsinkelsen vil være relativt små innenfor et gitt innspillingsrom (Everest, 2001, s. 429). Foregående setning er for så vidt bare en alternativ framstilling av det vi har sett på over en del sider nå. Davis påviste videre at lydteknikeren i et konvensjonelt, ubehandlet kontrollrom ikke er i stand til å oppfatte tomrommet mellom den direkte lyden og refleksjonene. Årsaken er at tidlige refleksjoner som oppstår i kontrollrommet vil maskere denne forsinkelsen, og lydteknikeren vil dermed bli frarøvet viktig informasjon for sin vurdering av lyden. Denne kunnskapen avdekket behovet for å kontrollere de akustiske forholdene i kontrollrom.

Det er ulike metoder som benyttes i tilpasningen av kontrollrom. De ulike metodene baserer seg hovedsakelig på to ulike prinsipper.

1. Man gjør den fremre delen av kontrollrommet refleksjonsfri ved å benytte høyt absorberende materialer denne delen av rommet. De første tydelige refleksjonene som når lydteknikeren vil da stamme fra bakre vegg og man oppnår en større forsinkelse i kontrollrommet enn i innspillingsrommet.
2. Man skaper en refleksjonsfri sone rundt lydteknikeren ved å konstruere fremre del av rommet på en slik måte at refleksjoner sendes i andre retninger enn mot lydteknikeren.

Ved begge teknikker benytter man gjerne overflater som gir høy grad av diffusjon i bakre del av rommet. Vi ser her et annet område som skiller det profesjonelle lydstudio fra hjemmestudioet. Det er antakeligvis også den viktigste årsaken til at man i profesjonelle produksjoner godtar grunnopptak i hvilke omgivelser som helst, men gjerne flytter miksing inn i kontrollerte omgivelser.

### **3.2 Effektprosessering/miksing**

Mikseprosessen er en meget viktig fase i tilblivelsen til en populærmusikkproduksjon. Nedmiksing av en innspilling innebærer en rekke estetiske valg som på mange vis har betydning for vår opplevelse av det ferdige produktet. Mikseprosessen blir ekstra interessant i

lyset av tema for denne oppgaven, da det er under nedmiksing man fastsetter det tredimensjonale lydbildet vi forholder oss til i reproduksjonen.

Profesjonelle lydstudioer har i dag tilnærmet full automasjon av det tekniske utstyret. Det innebærer blant annet at alle nivåer og innstillinger på miksebord og annet teknisk utstyr kan lagres, og dermed også hentes fram igjen ved senere anledning. Mulighetene som ligger i full automasjon har hatt den konsekvens at innspillingsfasen og miksefasen tenderer til å flyte mer sammen. Altså at man gjør en del nedmiksingsvalg samtidig som man foretar selve innspillingen. Selv om grensene mellom de to fasene er noe mer flytende i dag enn tidligere, er det allikevel både matnyttig og relevant å se på de to fasene separat.

Noen premisser vil selvsagt være lagt i det innspilte materialet, og ting kan også endres i løpet av etterproduksjon, men det er primært i nedmiksing man fastsetter spatiale trekk. Mulighetene til manipulasjon av det innspilte materialet er enorme. Mulighetene har nesten alltid vært store, men den digitale revolusjonen åpnet for en helt ny sfære av muligheter.

Grovt sett kan man dele de tekniske grep som utføres under nedmiksing inn i følgende:

- Manuell dynamikkontroll (avspillingsvolum)
- Automatisk dynamikkontroll (kompresjon/expandere/gater)
- Filtrering (EQ)
- Panorering (plassering på breddeaksen)
- Tilføring av romklang/effekter
- Digital redigering

En stor andel produksjoner begrenser seg til de grep som står i listen ovenfor, men mangfoldet er stort. Som nevnt ovenfor, kan innspillingsfasen og miksefasen til tider flyte sammen. I enkelte stilarter vil det jeg benevner som innspillingsfasen, også være en komposisjonsfase. Digitale redigeringsverktøy med sine ikke-lineære, grafiske grensesnitt, åpner for strukturelle og kompositoriske endringer når som helst i prosessen. Man må anta at mulighetene utnyttes.

De ulike tekniske grepene er tett forbundet med hverandre. Det er en gjensidig påvirkning mellom dem, og alle spiller en viktig rolle i konstruksjonen av det endelige lydbildet. I fortsettelsen skal jeg forsøke å se på romklangen separat, men uttalelsene må nødvendigvis ses i lys av det innbyrdes forholdet som eksisterer.

Fram til nå har jeg hatt fokus på de tidlige refleksjonene. Det er det fortsatt meningen å ha, men det er sannsynlig at mange lydteknikere og produsenter neppe opererer med en like

klar distinksjon. De sene refleksjonene vil derfor bli tatt litt mer med i betraktningen i det følgende.

De tidlige refleksjonenes rolle i mikseprosessen, utover lytteforholdene i kontrollrommet, begrenser seg til kunstige romklang i en eller annen form. I begrepsavklaringen i kapittel 1.7 så vi at kunstig romklang defineres som: *Romklang som er initiert ved reproduksjon av lydkilden, eller som manipuleres uavhengig av den direkte lyden.*” Hele definisjonen kan synes noe unødvendig. En definisjon ut fra praktisk tilnærming i lydstudio kan muligens også holde. Det er først og fremst bruken av impulsresponser ved konvulsjon, samt andre digitale teknikker, som fremtvinger definisjonen, men den tar også høyde for gamle tiders bruk av ekkorom. Både impulsresponser og ekkorom har gjerne et naturlig opphav, men det karakteristiske er at romklangen er initiert av en lydkilde som ikke nødvendigvis har samme sprednings- og frekvenskarakteristikk som lydkilden romklangen tilføres. Dette aspektet vil bli utdypet videre litt senere i kapittelet.

Populærmusikken benytter en rekke teknikker for å tilføre kunstig romklang til lydkilder. Hver teknikk har sitt eget særpreg, med sine fordeler og ulemper. Det karakteristiske ved bruken, uansett teknikk, er at man har separat kontroll over romklangen uavhengig av den opprinnelige lyden fra lydkilden.

Det er ikke uvanlig å benytte kunstig romklang også i løpet av selve innspillingen av en populærmusikkproduksjon, blant annet i hodetelefonene til musikerne, men det er sjelden man fanger opp og lagrer den kunstige romklangen som benyttes under innspillingen. Endelige valg av mengde og type romklang gjøres normalt i et mer helhetlig lys i løpet av mikseprosessen.

### **3.3 Kunstig romklang**

Det klassiske skillet mellom analog og digital teknologi i musikkverden gjelder også innenfor kunstig romklang. Historien startet med analoge teknikker, mens den digitale teknologien ble tilgjengelig sent på 70-tallet. Med den digitale teknikken søkte man både å gjenskape klangen av de analoge teknikkene, men samtidig også å skape nye innfallsvinkler for å simulere naturlig romklang. Både analoge og digitale teknikker benyttes fremdeles, men digital prosessering er nærmest enerådende. Analog klangplater har definitivt fått sin renessanse de siste årene, men de er forbeholdt de få.

I år 2006 benytter man i praksis to ulike teknikker for å generere kunstig romklang. Den første og, i praktisk anvendelse, eldste teknikken er en algoritmebasert metode. Ved sinnrike kombinasjoner av signalforsinkelser og rekursive filtre, er man i stand til å simulere

både naturlig og unaturlig klingende romklang. Beskrivelsen er selvsagt en forenkling, og hemmelighetskremmeriet rundt teknologien er stor, men Schroeders arbeid fra 1962 er fremdeles grunnlaget for denne teknikken (Griesinger, 1989, s. 5; Gardner, 1997).

Den andre teknikken har utgangspunkt i bruken av impulsresponser ved konvulsjon. En impulsrespons er kort sagt en kartlegging av et objekts reaksjon på en impuls. Det klassiske eksempelet på en impulsrespons får man ved å ta opp lyden av en oppblåst ballong som sprekker inne i en konsertsal. En ballong som sprekker er i praksis en meget kort impuls som vil generere lik energi over hele det hørbare frekvensregisteret, og i alle retninger. Selve impulsen varer kun noen få millisekunder, men i kjølvannet av impulsen vil man også høre de refleksjonene som oppstår i rommet som følge av impulsen. Et slikt lydopptak vil dermed være en kartlegging av romklangen i konsertsalen. Slike impulsresponser kan senere benyttes til å tilføre romklang til innspilte lydkilder ved konvulsjon. Den enkle beskrivelsen av konvulsjon er at man digitalt bytter ut ballongen med lydkilden.

En stor andel av de impulsresponser som benyttes til simulering av romklang er fanget inn i virkelige i rom, men man kan i teorien fange inn impulsen til hva som helst. To relevante eksempler er for eksempel analoge klangplater og algoritmebaserte romklangsimulatorer.

Man har hatt kjennskap til bruken av impulsresponser og konvulsjon lenge. Det er allikevel bare 10 år siden man begynte å se verktøy til praktisk anvendelse i lydstudioer. Årsaken ligger i at romklang ved konvulsjon er en regneintensiv oppgave. I lydstudioer ønsker man gjerne å arbeide i sanntid, og det innebærer at man må ha en del prosessorkraft tilgjengelig for å generere romklang ved konvulsjon. Det var i så måte utviklingen i tilgjengelig prosessorkraft på markedet som omsider åpnet for allmenn bruk av teknikken.

Den omtrentlige dobling av antall transistorer i prosessorer hver 18. måned, kjent som "Moore's lov" fra 1965 ("Wikipedia", 2006), har hatt betydning også for algoritmebasert romklang. Algoritmebasert romklang av høy kvalitet er også en regneintensiv oppgave. Tilgjengelig prosessorkraft har dermed alltid vært en viktig rammebetingelse for kvaliteten, som i så måte har avspeilt seg i pris. Prosessorkraft er dog ikke det eneste kriterium for kvaliteten. Det faktum blir meget tydelig når man ser at enkelte modeller fra midten av 80-tallet fremdeles er mye brukt og meget attraktive i markedet.

### **3.3.1 Naturlig og unaturlig romklang**

Hvilken av de to teknikkene – algoritmer eller impulsresponser – som er best egnet for tilføring av romklang i populærmusikkproduksjoner er en stor diskusjon. Man skulle anta at impulsresponser, med sine naturtro gjenskapninger av naturlige forhold, vil være

algoritmeklangene overlegne. Men det er ikke nødvendigvis slik at det helt naturlige alltid er det mest attraktive.

Man har i og for seg alltid visst at den beste romklangen til stereo reproduksjon ikke nødvendigvis er lik den beste romklangen for konsertsaler (Griesinger, 1989, s. 6), men det synes å ha skjedd et lite skifte i tilnærmingen til algoritmebasert romklang. Skiftet skjedde samtidig med impulsresponsenes inntog på romklangsulatormarkedet. Der det tidligere var et uttalt mål å best mulig gjenskape det naturlige (Griesinger, 1989, s. 1), er det naturlige nå en begrensende faktor for det estetiske uttrykket (Lexicon, 2000). Dette synet frontes naturlig nok av produsenter som har spesialisert seg på algoritmebasert romklang. Et raskt søk gjennom brukermanualene til romklangprodusenten Lexicons nyeste og foregående flaggskip, viser for eksempel at ordet ”natural” brukes 7 ganger i omtalen av nyeste modell, mot 24 ganger om den eldste (Lexicon, 2001; Lexicon, 1993).

Til tross for en til tider markedstilpasset retorikk er det flere argumenter som trekker i retning av at en algoritmebasert tilnærming har sine fordeler i musikkproduksjon. Argumentene kan i hovedsak koples til mulighetene for manipulasjon og kontroll, og dermed bedre muligheter for tilpasning av romklangens karakteristikk til det estetiske uttrykket man søker, men også den dynamiske karakteren til algoritmeklangene trekkes fram som en fordel. Impulsresponsen kan formidle et meget realistisk bilde av virkelige omgivelser, men det er i utgangspunktet et statisk bilde. I gode algoritmebaserte romklangsulatorer benytter man en grad av tilfeldighet i refleksjonene for å høyne kvaliteten på romklangen (Lexicon, 2001, s. 5-8). Denne tilfeldigheten korresponderer godt med romklangen som oppstår i naturlige omgivelser.

Noe av impulsresponsenes statiske preg kan kompenseres for ved funksjonalitet som muliggjør manipulasjon av enkelte faktorer, på lik linje med algoritmeklangene. Dette gjøres også av flere produsenter, med den retoriske konsekvens at applikasjonene blir mer algoritmepregede.

”The Plate sound is what most people associate with the word reverb, and it is useful for all popular music.” (Lexicon, 2001, s. 5-9)

Sitatet ovenfor synliggjør noe av den posisjonen analoge teknikker fremdeles har innen populærmusikkproduksjon, og samtidig at simulering av naturlige forhold ikke er det eneste kriteriet for bruken av kunstig romklang. Det skal sies at både algoritmer og impulsresponsen gjenskaper plateklanger meget bra.

Jeg skal nå gå litt nærmere inn på romklangen som romklangsimulatorer genererer. Jeg starter med den algoritmebaserte romklangen, for deretter å se på impulsresponsklange.

### 3.3.2 Algoritmebaserte klanger

Før impulsresponsteknologien kom på markedet var tilgangen på kunstig romklang av høy kvalitet forbeholdt profesjonelle lydstudioene. Det har siden 80-tallet eksistert romklangsimulatorer i alle prisklasser, men forholdet mellom pris og kvalitet har hele tiden gått hånd i hånd. Det har alltid vært, og synes fremdeles å være, bare noen få som klarer å produsere algoritmebasert romklang av virkelig høy kvalitet. Tidligere innebar det en tilnærmet monopolsituasjon, med tilsvarende høye priser. Selv om produsentene nå presses av impulsresponsmarkedet på enkelte områder, holder prisene seg fortsatt relativt høye.

I likhet med *akademia*, har romklangprodusentenes fokus beveget seg fra stereo til surround-lyd de senere årene. Flaggskipene til produsenter som Lexicon og TCElectronic er til dels utviklet med fokus på best mulig å formidle spatiale trekk i lyd reproduisert i 5.1 og 7.1 surround. Surroundbegrepet bør være kjent for de fleste, men for ordens skyld: 5.1 og 7.1-surround innebærer reproduksjon av lyd i henholdsvis 6 og 8 kanaler (les: høyttalere). Høyttalerne er plassert både foran og bak lytteren, og er spesielt utbredt i filmmusikk. At de nyere flaggskipene er utviklet med sikte på surround-lyd, innebærer absolutt ikke at modellene ikke fungerer til stereoproduksjon, men trenden synes til en viss grad å ha påvirket hvilke parametere brukeren kan kontrollere.

Fram til for 10 år siden ble nesten all kunstig romklang i populærmusikkproduksjon tilført ved bruk av såkalt generisk romklang. Begrepet er lånt fra TCElectronics brukermanualer (TCElectronic, 2005, s. 53). At den er generisk innebærer at samme romklang er tilført flere lydkilder, som igjen innebærer at romklangens karakteristikk i mindre grad er tilpasset den enkelte lydkildes plassering i lydbildet. Det motsatte er kildebasert romklang. Den teknologiske utviklingen, nok en gang som følge av Moores lov, har medført at man i dag har større mulighet til å behandle lydkilder uavhengig av hverandre. Muligheten for individuell behandling av lydkilder åpner for enda større grad av kontroll, men det er usikkert i hvilken grad det benyttes i populærmusikkproduksjoner per i dag. Teknologi til individuell behandling er fremdeles kostbar, og man må anta at den generiske romklangen vil være dominerende i noe tid framover.

#### *Parametere*

Vitenskapen rundt algoritmebasert modellering har alltid operert med et distinkt skille mellom tidlige og sene refleksjoner (Moorer, 1979). Noe av alibiet for denne oppgavens fokus på



tidlige refleksjoner ligger akkurat her. Distinksjonen har som vi allerede har sett, opphav i akustikkens tilnærming til romklang, men også de teknikker man benytter i modellering av romklang ved digital signalprosessering underbygger dette skillet.

Det karakteristiske ved tidlige refleksjoner kontra sene refleksjoner er at førstnevnte gruppe er relativt få i antall sammenlignet med sistnevnte. Vi har også sett at den enkelte tidlige refleksjon, avhengig av ankomsttidspunkt, frekvenskarakteristikk og relasjon til andre refleksjoner, kan ha stor betydning for vår opplevelse av lyden. De sene refleksjonene på sin side påvirker hovedsakelig vår opplevelse i kraft av den gruppen de til sammen utgjør. Denne grunnleggende forskjellen gjenspeiles som nevnt i måten man modellerer refleksjonene på, og dermed også de parameterne man gis mulighet til å kontrollere i den kunstige romklangen.

Den enkelte produsent av romklangsimulatorer har naturlig nok sin egen tilnærming i utformingen av brukergrensesnittene. Enkelte grunnleggende parametere benyttes av alle, mens andre aspekter angripes noe ulikt mellom ulike modeller og produsenter. Parameterne kan grovt sett deles inn i følgende grupper:

- Etterklangtid (RT60), inklusiv frekvensavhengig etterklangtid
- Rommets størrelse og form
- Rommets frekvenskarakteristikk, grad av diffusjon og hastighet på oppbygging av refleksjoner
- Initiale forsinkelser, ankomsttidspunkt for første tidlige refleksjon og sene refleksjon
- Karakteristikk og mengde tidlige refleksjoner
- Nivåforhold mellom direkte lyd, tidlige refleksjoner og sene refleksjoner

Hvor detaljert man kan kontrollere de ulike aspektene varierer fra modell til modell. Hvis vi holder oss til tidlige refleksjoner gir enkelte modeller brukeren muligheten til presis kontroll av den enkelte initiale refleksjon, så vel som den overordnede karakteristikken og spredningen av de tidlige refleksjonene. Andre modeller starter med forhåndsdefinerte kategorier, gjerne med utgangspunkt i grunnleggende romtyper med parametere som bestemmer den overordnede karakteristikken.

De fleste modeller kommer med et stort antall ferdige programmer som simulerer mange ulike typer rom. Den jevne bruker av produktene vil gjerne ta utgangspunkt i et av de mange programmene, for deretter å justere parametere etter behov. De enorme variasjonsmulighetene i dagens romklangsimulatorer genereres normalt ut fra noen få grunnleggende algoritmer i hver modell.

### *Kvalitet*

Fordi de tidlige refleksjonene er relativt få i antall og derfor ganske forutsigbare, skulle man tro at simuleringen av disse ville være forholdsvis enkel. Man skulle derfor også tro at kvaliteten på romklangen i hovedsak ble bestemt av karakteristikken til de sene refleksjonene. Det er ikke nødvendigvis slik. Karakteristikken til de sene refleksjonene er definitivt avgjørende, og sannsynligvis det mest umiddelbare kriteriet for en vurdering av kvaliteten. Samtidig synes uttalelser fra en rekke lydteknikere å peke i retning av at deres preferanser i stor grad styres av aspekter som må knyttes til de tidlige refleksjonene. ”*For meg har det med hvor godt jeg synes klangene blander med resten av lyden.*” (Engine Ears, 2002) Sitatet går i favør av produsenten Lexicons klanger, og føyer seg inn i en rekke av lignende uttalelser undertegnede har hørt/lest fra ulike lydteknikere. Lexicons evne til å simulere diffusjon også i de tidlige refleksjonene har nok en del av æren for produsentens særstilling gjennom mange år (Geoffrey, 2001, s. 1.87).

### *Refleksjonenes karakteristik*

Vi har tidligere sett at de tidlige refleksjonenes karakteristik påvirkes av en rekke faktorer, deriblant lydildens posisjon i rommet. Vi vet også at refleksjonenes karakteristik har stor betydning for hørselens bedømming av lydildens posisjon og framtoning. Det blir dermed et relevant spørsmål hvorvidt man er i stand til å gjenskape dette mangfoldet i kunstig generert romklang. Det enkle svaret på spørsmålet er: Ja, innenfor de rammebetingelser som legges av stereomediet. Med andre ord, man er i minst like god stand til å generere tidlige refleksjoner som mikrofoner er i stand til å formidle de naturlige. Det naturlige oppfølgingsspørsmålet til dette blir: Gjør man dette i populærmusikkproduksjoner? Det enkle svaret denne gang er: Nei, stort sett ikke.

Årsaken er å finne i den utstrakte bruken av generisk romklang. Man må som tidligere nevnt anta at bruken av kildebasert romklang har økt betraktelig de siste 10 årene. Den vil sannsynligvis fortsette å øke parallelt med den teknologiske utviklingen, men generisk romklang vil antakeligvis ha en markant rolle i god tid framover.

Signalgangen ved tilføring av kunstig romklang kan deles inn i to ulike teknikker: mono-til-stereo og stereo-til-stereo. Det første ordet beskriver signalet man sender inn, mens det andre beskriver signalet som kommer ut av romklangsformatoren. Signalgangen legger åpenbare premisser for refleksjonene som genereres. Dersom signalet som sendes inn er mono, vil man teoretisk bare kunne generere tidlige refleksjoner som simulerer den presise posisjonen til en enkelt lydild. Er inngangssignalet stereo, kan antallet lydild økes til to. Et signal som bare består av henholdsvis en eller to lydild kan anses å være kildebasert

romklang, forutsatt at romklangsimulatoren er i stand til å simulere presise posisjoner. En stor andel av modellene på markedet er ikke i stand til å gjøre dette.

Det karakteristiske ved generisk romklang er at signalet som sendes inn i romklangsimulatoren gjerne inneholder en miks av flere lydkilder. Dersom inngangssignalet er i mono vil alle lydkildene bli tilført romklang med identisk karakteristikk. Med identisk karakteristikk menes likt ankomsttidspunkt etter den direkte lyden, samt innbyrdes lik distribusjon, diffusjon, amplitude og frekvenskarakteristikk på refleksjonene. Man kan naturlig nok regulere D/R-ratioen individuelt, men utover det vil romklangens preg kun variere med lydkildens egen karakteristikk. Forkortelsen D/R-ratio er engelsk og står for "Direct to Reverberant ratio". Forkortelsen peker på nivåforholdet (ratioen) mellom direkte lyd og reflektert lyd.

Er inngangssignalet i stereo vil karakteristikken kunne variere noe mer, forutsatt at lydkildene er ulikt panorert. Den overordnede karakteristikken vil fremdeles være lik, men på grunn av muligheten for panorering vil det oppstå større variasjoner avhengig av lydkildenes posisjon i signalet. Man kan sammenligne det som skjer med de refleksjonene som oppstår i et lytterom når man lytter til to høyttalere i stereo. Stereobildet fra høyttalerne vil kunne reprodusere en rekke lydkilder med ulik plassering, men refleksjonene som oppstår i lytterrommet blir til ut fra to spesifikke punkter.

Generisk romklang setter klare premisser for modelleringen av de tidlige refleksjonene. I historisk perspektiv er nok dette å snu litt på rollene, men i det man tilfører samme romklang til flere lydkilder er det begrenset hvor presist man ønsker de tidlige refleksjonene skal simulere eksakte posisjoner.

Foregående avsnitt kan muligens skape et inntrykk av at de tidlige refleksjonene i generisk romklang ikke vil ha noen innvirkning på lydkildenes spatiale trekk. Det stemmer overhode ikke. Går man tilbake til kapittel 2 og ser på detaljene rundt de spatiale trekkene tidlige refleksjoner påvirker opplevelsen av, ser vi at ingen automatisk diskvalifiseres av generisk romklang. Det er mer snakk om at den generiske romklangen har en redusert evne til å formidle spatiale trekk sammenlignet med den kildebaserte romklangen. Forklaringen på den reduserte evnen er å finne i at alle lydkilder som tilføres generisk romklang vil generere refleksjoner med lik karakteristikk. I det flere lydkilder genererer refleksjoner med lik karakteristikk, må man anta at det oppstår ulike grader av maskering av refleksjonene til den enkelte lydkilde. Kort sagt, refleksjonene til den enkelte lydkilde blir ikke så tydelige som de kunne vært. Konsekvensen er en dårligere evne til å formidle spatiale trekk.

Omtalen av de kunstig genererte refleksjonene har så langt fokusert på betydningen for framtoningen til den enkelte lydkilde. Man kan fort komme i skade for å glemme det opplagte, at romklangen også formidler romlighet. Tilføring av generisk romklang innebærer at man skaper en romlig sammenheng mellom flere lydkilder. Dette vil bli behandlet lenger nede, men først skal vi se på impulsresponsen.

### **3.3.3 Impulsresponsen**

Romklang ved konvulsjon har vært en stor gave til amatører og den halvprofesjonelle. Ved å gjøre tilgangen til romklang av høy kvalitet langt rimeligere har metoden redusert avstanden mellom hjemmestudioet og det profesjonelle studioet. Kvaliteten er definitivt meget bra. Ankepunktet mot impulsresponsen er at de i utgangspunktet er helt statiske. Ulike produsenter har ulike metoder for å gjøre bruken mer dynamisk. Teknologien er fremdeles relativt ung og må derfor anses å være i en utviklingsfase. Vi skal nå se på noen av de utfordringene som er knyttet til impulsresponsens statiske preg.

En konvulsjon av en innspilt lydkilde og impulsresponsen til et rom blir gjerne sidestilt med om man skulle tatt opp lydkilden i rommet impulsresponsen er hentet fra. Det stemmer ikke helt. Det er flere grunner til det. Den viktigste årsaken er at lydkilden (impulsen) som initierer impulsresponsen har en annen spredningskarakteristikk enn lydkilden man senere behandler ved konvulsjon.

Der man tradisjonelt har stukket hull på ballonger og fyrt av pistolskudd for å initiere impulsresponsen, benytter man i dag gjerne høyttalere til å generere lyden som danner utgangspunktet for impulsresponsen. Det benyttes ulike teknikker, alle med sine fordeler og ulemper, men grovt sett ender man opp med samme resultat. Det karakteristiske ved høyttalere som benyttes i seriøse opptak av impulsresponsen er at de har en jevn kuleformet spredningskarakteristikk. De fleste instrumenter har langt fra en like jevn spredningskarakteristikk. For et instrument som er tilført romklang ved konvulsjon, vil ulikhetene i spredningskarakteristikk innebære at noen refleksjoner vil framstå kraftigere enn de ville gjort om det var instrumentet selv som ble tatt opp i rommet der impulsresponsen er hentet fra.

Et annet aspekt som gjør at romklang ved konvulsjon ikke er en tro kopi av en naturlig opptakssituasjon, oppstår dersom lydkilden er tatt opp med mikrofon. Forutsatt at den ikke er spilt inn i ekkofrie omgivelser vil det allerede være refleksjoner til stede i det opprinnelige signalet. De opprinnelige refleksjonene vil naturlig nok også følge med videre og vil potensielt prege den tilførte romklangens karakteristikk.

Foregående avsitt er selvsagt ikke unikt for bruken av impulsrespons. Det gjelder ved tilføring av enhver form for kunstig romklang. I hvilken grad de opprinnelig innspilte refleksjonene vil prege det endelige resultatet, avhenger i særdeleshet av nivåforholdet mellom direkte lyd og refleksjoner, men også refleksjonenes generelle karakteristikk.

Det er også diskutabelt i hvilken grad avviket fra det helt naturlige vil prege opplevelsen. Som vi har sett i omtalen av algoritmebasert romklang er det ikke slik at et avvik fra det naturlige nødvendigvis er uønsket. Avsnittene ovenfor må i så måte ikke leses som en vurdering av romklangens estetiske kvalitet i seg selv, men belyser for så vidt en svakhet dersom en gjenskaping av det helt naturlige er målet.

Den neste utfordringen knyttet til impulsresponsenes statiske preg oppstår når man benytter impulsresponsen til generisk romklang. Impulsresponsens karakteristikk gjenspeiler en meget presis posisjon. Ved kildebasert romklang kan det være en fordel, men det er som vi har sett lenger oppe heller en ulempe ved generisk romklang.

Bruken av impulsrespons åpner for god simulering av flere eksakte posisjoner innenfor samme rom, for eksempel til bruk i kildebasert romklang, men det forutsetter at man på forhånd har fanget inn impulsrespons som er initiert fra ulike posisjoner i rommet. Impulsrespons er i utgangspunktet lite fleksible sammenlignet med den algoritmebaserte romklangen. Grunnleggende parametere i bruken av kunstig romklang, slik som etterklangstid (RT60) og initial forsinkelse (predelay), er allerede fastsatt. Dette er lite ønskelig i populærmusikkproduksjon. Lydteknikeren kan selvsagt lete opp impulsrespons som har akkurat den karakteristikken produsenten ønsker, men det er en relativt upraktisk metode. For å øke anvendeligheten til den enkelte impulsrespons åpner de fleste applikasjoner derfor for en parameterstyrt manipulasjon av impulsresponsens karakteristikk. Manipulasjonen tar gjerne utgangspunkt i en digital analyse som søker å avdekke impulsresponsens bestanddeler. Den initiale analysen kan eventuelt være assistert av indikasjoner fra brukeren. Når brukeren så endrer på et eller flere parametere, genererer applikasjonen en ny impulsrespons på bakgrunn av de endringene brukeren har gjort. Det er stor variasjon i hvor vellykket denne form for manipulasjon er. Noen parametere, slik som initial forsinkelse og overordnet frekvenskarakteristikk, lar seg langt bedre manipulere enn andre parametere.

### **3.4 Bruken i praksis**

Jeg vil avslutte dette kapitlet med et noen avsnitt som belyser lydteknikere og produsenters praktiske tilnærming til bruk av kunstig romklang. Doktrinen om at ”alt er lov i populærmusikkproduksjon” gjelder også her. Beskrivelsene er derfor ikke annet enn noen

ytterst få eksempler fra et mangfold som er enormt. Jeg tror allikevel eksemplene vil være nyttige for leseren, til tross det begrensede antallet eksempler.

Tilføring av kunstig romklang i en lydmiiks, være seg til en eller flere lydkilder, generisk eller kildebasert, skjer normalt i et helhetlig perspektiv. Romklangen kan og vil gjerne fungere som et musikalsk element, men det er antakeligvis som en formidler av spatiale trekk i lydbildet vi kjenner den best. I arbeidet med konstruksjonen av lydbildet er romklangen ett av flere verktøy lydteknikeren har til rådighet. Denne oppgavens løsrivelse av romklangen gjenspeiler i så måte ikke lydteknikerens mer helhetlige innfallsvinkel. I beskrivelser som omhandler lydteknikerens praktiske arbeid er det nødvendigvis vanskeligere å løsrive romklangen på samme måte.

Et annet aspekt som også avviker fra denne oppgavens tilnærming til emnet er at langt fra alle lydteknikere skiller tidlige og sene refleksjoner like klart fra hverandre. Man kan muligens tro at dette gjelder lydteknikere som mangler kunnskap, men det synes absolutt ikke å stemme. Vi skal se et eksempel på dette litt lenger nede.

#### *Visualisering av lydbildet*

En beskrivelse som synes å gå igjen hos mange lydteknikere er at de visualiserer et band på en scene når de skal plassere instrumentene i lydbildets tre dimensjoner. Med ulike tekniske grep, deriblant romklang, plasseres instrumentene i ulike posisjoner på scenen.

#### *Flere typer romklang*

”You may hear two or more reverb settings used on a typical pop record, often one setting for the vocals and another for the drums and percussion.” (White, 2001b, s. 1)

Bruk av flere typer romklang i en og samme miiks er meget utbredt. I opplæringslitteratur knyttet til emnet finner man også gjerne generelle anbefalinger som ”plateklang til perkussive instrumenter” og ”hall til vokal”. En viktig funksjon til romklangen er gjerne å binde de ulike instrumentene i innspillingen bedre sammen. Til dette arbeidet er det ikke uvanlig å benytte også en tredje type romklang. Følgende sitat illustrerer meget godt en variant av dette temaet:

”Når jeg mikser, er låtens fulle rom en kombinasjon av tre klanger. Et bittelite rom, en medium stage med pre delay på omkring 30 ms, og tilslutt en medium hall med pre daly opp til 80 ms. De forskjellige instrumentene får en kombinasjon av en eller flere av disse klanger. Da har jeg mye mere kontroll på plassering i rommet. Noe trenger bare størrelse, ikke hale, det får det bittelille rommet, noe trenger hale ikke størrelse, det får kun hallen.” (Engine Ears, 2002)

Forfatteren av sitatet er administrator på nettstedet til firmaet Engine Ears. Undertegnede har ikke forsøkt å identifisere forfatteren ytterligere, men beskrivelsen korresponderer i høy grad med lydtekniker og produsent Truls Birkelands beskrivelse av egen arbeidsmetode da

undertegnede intervjuet ham under arbeidet med hovedoppgaven i 1998. Truls Birkeland er ansatt i Engine Ears.

## 4 Konsumentens lyttesituasjon

Konsumenters lyttesituasjon er alt annet enn ensartet. Vi mennesker lytter til reprodusert musikk i et stort mangfold av rom og ulike sosiale sammenhenger. Intensjonen med denne delen er ikke å avdekke alle akustiske/psykoakustiske aspekter knyttet til ulike lyttesituasjoner, til det er både emnet og mangfoldet for stort. Men ved å peke på noen viktige aspekter i konsumenters lyttesituasjon, håper jeg å formidle et bilde av dette feltets betydning for lytteopplevelsen.

Hovedfokus blir på noen av de lyttesituasjoner som ofte oppstår innenfor hjemmets fire vegger. Andre områder vil også bli berørt, men da mer overflatisk. Fokus blir ikke bare begrenset til romklang. Aspekter knyttet til lyttesituasjonens betydning for andre spatiale trekk vil også bli tatt med.

Den optimale lyttesituasjonen må nødvendigvis være målestokken for en vurdering av den ikke-optimale lyttesituasjonen. Målsettingen for denne delen blir i så måte å peke på steder hvor den ikke-optimale lyttesituasjonen avviker fra den optimale, og i forlengelsen av det peke på hvordan disse avvikene arter seg.

En stor utfordring knyttet til det å skulle beskrive konsumentens lyttesituasjon ved lytting til reprodusert musikk er at det er utført lite forskning på området. Det er muligens en sterk forenkling, men hovedinntrykket er at forskningen relativt enkelt slår fast at det oppstår en forringing av det optimale. Få synes å ha tatt tak i hvordan forringelsen arter seg for lytteren.



## 4.1 Effekten av ulike lyttesituasjoner

En grei plass å starte er å ta tak i hva som skjer i det vi begynner å bevege oss bort fra sweet spot. For den bevisste lytter vil effekten av en slik bevegelse fortone seg som en umiddelbar fordreining og/eller innsnevring av lydbildets spatiale trekk. Størrelsen på sweet spot kan for øvrig variere en hel del mellom ulike oppsett, også i bredden.

Generelt kan man si at bevegelser hvor man ikke opprettholder symmetrien mellom høyttalerne er mest utsatt for endring av stereobildet. Bevegelse langs aksene som strekker seg fra midt mellom høyttalerne og gjennom sweet spot vil i teorien kunne opprettholde et presist stereoperspektiv, men lydbildets detaljrikdom snevres ubønnhørlig inn med økende avstand.

Ved bevegelse bort fra midtaksen vil det oppstå en forskyving av symmetrien mellom lytterens to ører og signalene fra høyttalerne. I det symmetrien forrykkes vil ikke relasjonene mellom fase (ankomsttidspunkt) og amplitude i den kombinerte lyden fra høyttalerne lenger være presise og hørselen vår vil tolke summen av signalene i en annen retning. Dette har ulike konsekvenser. I retningsbestemmelsen til en lydkilde i den reproduserte lyden vil det i starten vanligvis oppleves som en forskyving av lydkildens posisjon. I opplevelsen av rom innebærer det vanligvis en forflating og reduksjon i de romlige attributter hørselen klarer å tolke ut av lydbildet.

Forutsatt at vi beveger oss bort fra høyttalerne vil de akustiske forholdene i lytterrommet spille en gradvis større rolle i lyden vi opplever. Ved økt avstand til høyttalerne, endrer vi nivåforholdet mellom den direkte lyden og refleksjonene fra rommet, den såkalte D/R-ratioen. En lavere D/R-ratio (økt avstand) innebærer generelt at avspillingsrommets akustikk får en større rolle i vår opplevelse av den reproduserte lyden, og vi får en degradering i opplevelsen av spatiale trekk.

Den observante leser vil legge merke til at ord som ”generelt” og ”vanligvis” benyttes i utstrakt grad. Årsaken er at det er få aspekter som kan sies å være helt allmenngyldige i dette feltet. Det hersker fremdeles uenighet rundt flere aspekter knyttet til reproduksjon i stereo, og på flere områder mangler det kunnskap. Til og med rundt basale aspekter som viktigheten av symmetrien mellom lytterens ører og høyttalerne råder er det til dels uenighet forskere i mellom (Rumsey, 2001, s. 57). Litt av årsaken til akkurat det er at selv om sweet spot er et relativt snevert område, er høresansen noe tilgivende med hensyn til hodebevegelser. Eventuelle hodebevegelser forrykker også symmetrien mellom ører og høyttalere. Helt presis reproduksjon forutsetter allikevel et hode som er vendt forover mot høyttalerne (Griesinger, 2000, s. 14).

En viktig årsak til den utstrakte bruken av ordene ”generelt” og ”vanligvis” er at det er for store variasjoner mellom resultatene av ulike oppsett av høyttalere og lytterrom til at man kan generalisere helt. Som et eksempel til foregående setning, hevder Manny LaCarrubba ved Moulton Laboratories å ha konstruert et eksepsjonelt kontrollrom med en ganske annerledes tilnærming til interaksjonen mellom høyttalere og tidlige refleksjoner fra rommet enn en konvensjonell tilnærming skulle tilsi (LaCarubba, 1999). Det relative preget til tross, er det allikevel mer enn sannsynlig at det som beskrives i kapittelet er allmenngyldig for den store majoriteten av konsumenter.

Som vi har sett i kapittel 2 om tidlige refleksjoner vil et økt nivå og endret ankomsttidspunkt for de tidlige refleksjonene kunne innebære en lang rekke perseptuelle endringer. Det samme gjelder naturlig nok også for et økt nivå av sene refleksjoner, men mens de tidlige refleksjonenes innvirkning varierer med lytterens posisjon, forblir de sene refleksjonenes karakter relativt lik mellom posisjoner, i alle fall i mindre rom. De sene refleksjonenes betydning bestemmes derfor hovedsakelig av refleksjonenes lydstyrke, samt frekvenskarakteristikk.

Den mest opplagte konsekvensen av en lavere D/R-ratio, er en høyere grad av maskering. Små nyanser og detaljer som er hørbare i sweet spot vil skjules som en konsekvens av kraftigere refleksjoner fra rommet. Maskeringen kan være en konsekvens av både tidlige- og sene refleksjoner. En lavere D/R-ratio vil generelt senke forståeligheten (*engelsk: intelligibility*).

De sene refleksjonene vil i hovedsak fungere som maskerende støy i disseksjonen av det reproduserte lydbildet, men de er samtidig en meget viktig meningsbærer i interaksjonen mellom musikk og lytterrom. Lytterens opplevelse av musikken må alltid ses i lys av omgivelsene lytteren befinner seg i. I det lyset er de sene refleksjonene meget avgjørende. Dette aspektet faller litt til side for fokus her, men det er allikevel så viktig at det bør nevnes.

Opplevelsen av innhylling må antas å oppstå relativt sjelden ved lytting til konsumentanlegg, men forutsatt at den er til stede forutsetter den normalt en lytteposisjon i sweet spot. Eksternalisering er nok opplevelsen som råder i konsumentenes musikkopplevelse.

## **4.2 Nærhet til vegger**

Nærhet til vegger og andre store overflater er et trekk som nok kjennetegner en betydelig andel av vanlige lytteposisjoner i konsumenters hjem. Årsaken er at vi vanligvis sitter når vi lytter til musikk, og i og med at både sofaer, skrivepult og stoler rundt bord gjerne plasseres

i nærheten av vegger, er det også her vi ofte befinner oss når vi lytter til musikk. I posisjoner nær store overflater vil påvirkningen av tidlige refleksjoner være maksimal. Det er også inne ved veggene forsterkningen av rommets resonanser er på sitt kraftigste.

Det er forsket en del på opplevelsen av lydkilder avhengig av lytterens posisjon i hverdagslige omgivelser. Shinn-Cunningham har som tidligere nevnt avdekt hvordan posisjoner tett inntil vegger, med det ene øret rettet inn mot veggen, kan gjøre det lettere å bedømme posisjonen til en lydkilde (Shinn-Cunningham, 2003, s. 5).

Shinn-Cunningham har også påvist at det i små, reflekterende rom, som baderom, oppstår spatial maskering ved økt avstand mellom lydkilder (Shinn-Cunningham, 2003, s. 6). Det normale er at det oppstår spatial avmaskering ved økt avstand. For oss er dette kanskje mest relevant den dagen man skal installere stereoanlegg på baderommet sitt for utelukkende å lytte til innspillinger i "split-mono" fra tidlig 60-tall. Men det illustrerer samtidig at ulike omgivelser har ulik innvirkning på lydopplevelsen, og at denne innvirkningen ikke er helt forutsigbar.

Den generelle forskningen på opplevelsen av lydkilder har naturlig nok gyldighet for både levende og reprodusert lyd, men overføringsverdien til opplevelsen av reproduksjon i stereo er allikevel noe begrenset. Årsaken ligger i at de innspilte lydkildene reproduseres av to lydkilder, som i kombinasjon skal utgjør en helhet. Mono lydkilder som er panorert helt til venstre eller høyre, vil naturligvis kun bli spilt av i en høyttaler, men utover det vil hørselens plassering av innspilte lydkilder være basert på summen av to signaler. Resultatet blir lett meget forskjellig fra det forskningen på enkle lydkilder tilsier.

## **4.3 Dimensjonene**

Reproduksjonen av spatiale trekk relaterer seg naturlig nok til de tre dimensjonene vi lever i. Jeg skal nå ta et blikk på hvilke muligheter spatiale trekk i de tre dimensjonene har til å overleve i posisjoner utenfor sweet spot.

### **4.3.1 Høyde**

Høydedimensjonen i reproduksjon i stereo kan i all hovedsak relateres til frekvenshøyde, hvor lave frekvenser oppleves lavt, mens høye frekvenser oppleves høyere oppe. Det finnes produksjonsteknikker som gjør det mulig å utnytte høydedimensjonen bedre, men generelt er frekvenshøyde den avgjørende faktoren.

Teknikker for bedre plassering i høydedimensjonen har vanligvis et utgangspunkt i hodets innvirkning på lyden som når ørene våre, såkalte hoderelaterte overføringsfunksjoner (*engelsk HRTF = Head Related Transfer Functions*). Kort forklart går det ut på at lyden som

når trommehinnene farges (les: filtreres) avhengig av lydens innfallsvinkel til hodet. Ved å kartlegge overføringsfunksjonen for ulike vinkler, kan man gjenskape retningen til lydkilden i binaural reproduksjon. Effekten er enklest å oppnå i hodetelefoner, men det finnes også teknikker som utnytter dette til reproduksjon i høyttalere. Teknikken er spesielt utbredt i simulering av surround lyd i moderne fjernsynsapparater, men benyttes også i enkelte populærmusikkproduksjoner

Det synes som HRTF-applikasjonene som finnes på markedet sliter en del med posisjonering i høydedimensjonen ved reproduksjon i konvensjonell stereo. Lydeksempelen på nettsiden til innstikkprodusenten WaveArts vitner om det (Wave Arts, 2006). Du kan laste ned lydeksempel for avspilling i hodetelefoner, men ikke for høyttalere.

Med utgangspunkt i at høydedimensjonen i hovedsak defineres av frekvensinnholdet vil man i "beste fall" snakke om relativt marginale endringer i opplevelsen. Dette fordi frekvensinnholdet overlever noenlunde likt over kortere avstander. Høyere frekvenser absorberes selvsagt raskere i sin ferd gjennom luften, men i de omgivelser vi ser på her vil det være neglisjerbart. Det er selvsagt noe annet om avstanden mellom høyttalere og lytter blir veldig lang, eller lyden på sin ferd må rundt avrundede hjørner eller gjennom obstruksjoner som dører og lignende.

#### **4.3.2 Breddde**

Breddedimensjonen er langt mer kompleks enn høydedimensjonen. Det er også den dimensjonen som er viet desidert mest oppmerksomhet i forskningen så langt.

Den sterkt forenklede versjonen av opplevd endring i breddedimensjonen, ved økt avstand til høyttalerne, er en innsnevring og eventuell vridning av hele lydbildet, samt en utsmøring av de enkelte instrumentenes plassering og utstrekning. Virkeligheten er langt mer komplisert. Emnet er for stort til å gå i dybden på, men to ulike eksempler kan illustrere noe av det som skjer.

I avsnittet om metode i kapittel 1 ble det nevnt at jeg i forberedelsene til oppgaven gjennomførte et forforsøk som i tillegg til uttesting av metode, tok sikte på å avdekke sider ved konsumentens evner og innfallsvinkel ved lytting til produsert musikk. Den ene oppgaven i ABX-delen av forsøket kan være med på å illustrere hvordan en posisjon til side for midtaksen, med det ene øret vendt inn mot veggen, endrer lydbildets framtoning. I den ene oppgaven, som var en innspilling med gitar og vokal, lå den eneste forskjellen mellom A- og B-eksempelet i panoreringen av gitaren. Vokalen var sentrert, mens gitaren var panorert noe til henholdsvis venstre og høyre for senter. Ved lytting til eksemplene ved spisestuebordet,

var undertegnedes subjektive observasjon at det ikke lenger var mulig å plassere gitaren til venstre eller høyre for vokalen, men det var fremdeles en distinkt forskjell mellom de to eksemplene. En fremdeles subjektiv beskrivelse av forskjellen er at opplevelsen kan sammenlignes med å vende fasen til en av høyttalerne når man sitter i sweet spot. Forsøkspersonene som gjennomførte forsøket utenfor sweet spot klarte ikke å skille de to eksemplene fra hverandre. Oppgaven ble for øvrig antatt å være relativt enkel å løse for deltakerne. Noe den absolutt ikke var, selv i sweet spot.

Det andre eksempelet er litt mer vitenskapelig fundert og er hentet fra en pilotundersøkelse med hensikt å belyse grafiske metoder til avdekking av spatiale trekk i reprodusert lyd. I forsøket plasserte Rumsey, Bruyn og Ford forsøkspersonene i og utenfor sweet spot (Rumsey, Bruyn, Ford, 2001). I forsøket ble det også benyttet tre ulike høyttaleroppsett, hvor to av høyttaleroppsettene ikke fulgte retningslinjene for optimal gjengivelse.

På flere områder stemte observasjonene overens med antagelsene. Blant annet var opplevelsen av lydbildets bredde (*engelsk: image width*) tydelig bredere ved optimal posisjon og optimalt høyttaleroppsett. (Rumsey, Bruyn, Ford, 2001, s. 6). På andre områder indikerte besvarelsene avvik fra det man forventet. Tabell 4-1 summerer noen interessante observasjoner i forsøket:

Spatialt trekk	Medvirkende faktor	Effekt
Lybildets bredde	Lytterens posisjon	Optimal lytteposisjon resulterer i opplevelsen av bredere lydbilder
	Lytterens posisjon	Større avstander til høyttalerne påvirker ikke lydbildets bredde som forventet
	Lytterens posisjon	Kildematerialet påvirker lydbildets bredde i optimal lytteposisjon
	Høyttalernes posisjon	Ikke-optimale høyttalerposisjon påvirker lydbildets bredde i kombinasjon med ikke-optimal lytteposisjon
	Lytterens posisjon	Optimal lytteposisjon resulterer i mindre enn ikke-optimale posisjoner
Vridning av lydbildet	Høyttalernes posisjon	Større vridning ved stor avstand mellom høyttalerne ( $\pm 60^\circ$ )

**Tabell 4-1:** Observasjoner i lytteforsøk utført av Rumsey, Bruyn og Ford hvor man varierte både lytteren og høyttalernes posisjon. Innholdet i tabellen er et utdrag fra tabell i Rumsey, Bruyn og Fords artikkel. Egen oversettelse (Rumsey, Bruyn, Ford, 2001 s. 6)

Det kanskje mest interessante i oppramsingen i tabell 4-1 er observasjonen av at økt avstand til høyttalerne ikke påvirket lydbildets bredde slik forfatteren hadde forventet det ville gjøre.

### 4.3.3 Dybde

På lik linje med de to andre dimensjonene vil opplevelsen av avstand og dybde i lydbilder variere med lytterens posisjon i forhold til høyttalerne. Høyttalernes plassering i seg selv er

naturlig nok også av stor betydning. Dette aspektet ble så vidt berørt i avsnittet om bredde, men vil bli grundigere behandlet lenger nede.

Vår opplevelse av dybde dimensjonen i lydbilder kan deles inn i tre faktorer:

1. Avstand til den enkelte lydkilde
2. Dybden på hele lydbildet, fra fremre kant til bakre kant
3. Eventuell dybdeutstrekning til lydkilder

For å begrense omfanget har jeg valgt å fokusere på avstand i denne delen. Valget av fokus er gjort i tråd med lytteforsøkets fokus på avstand.

Hørselen benytter et stort mangfold av teknikker i sin analyse av lydbølgene som treffer ørene våre. Analysen av ethvert lydbilde, reprodusert eller ikke, utføres med bakgrunn i den enorme erfaringen høresansen har bygget opp i løpet av livet. Naturskapte lydkilder, i naturlige omgivelser, bærer normalt med seg flere indikatorer som av hørselen tolkes i samme retning. For eksempel vil følgende tre aspekter indikere en lydkilde som nærmer seg lytteren:

1. Lydkilden øker gradvis i lydnivå
2. Lydkilden får gradvis mer høyfrekvent innhold
3. D/R-ratioen øker gradvis

Når hørselen opplever motstridende indikatorer, og det skjer gjerne når vi lytter til reprodusert lyd, synes den å operere på bakgrunn av en form for flertallsbasert logikk. Den sammenstiller informasjonen som er tilgjengelig, og konkluderer ut fra den tolkningen som er mest sannsynlig (Rumsey, 2001, s. 34). At dette har stor betydning for opplevelsen av populærmusikkproduksjoner, blir spesielt tydelig i når vi ser på avstandsbedømmingen.

Hvis vi ser på karakteristikken til en lydkilde nær en lytter, og sammenligner den med karakteristikken til den samme lydkilden opplevd lenger borte, vil forskjellene i sistnevnte bestå i følgende:

1. Lavere lydstyrke (på grunn av større reiselengde)
2. Mindre høyfrekvent innhold (på grunn av luftens absorpsjon)
3. Mer romklang (i reflekterende omgivelser)
4. Mindre forskjell mellom ankomsttidspunkt av direkte lyd og første gulvrefleksjon
5. Dempet bakkerefleksjon

(Rumsey, 2001, s. 35)

Med det lydtekniske mangfoldet som preger produksjonen av populærmusikk, sier det seg selv at det ikke alltid er samsvar mellom alle disse aspektene. I etterforskningen av opplevd avstand oppstår det i så måte en differensiering mellom opplevelsen av de lydkildene som opprettholder et samsvar og de som ikke gjør det. For sistnevnte gruppe blir det et spørsmål om hvilke indikatorer som dominerer over andre.

I denne sammenheng er det mest aktuelt å se på forholdet mellom de 3 første indikatorene. Alle er i kraft av seg selv i stand til formidle en endring i opplevd avstand. Hvilken som er den dominante av de tre, varierer veldig avhengig av ulikt kildemateriale. Et generelt trekk i opplevelsen av lyd er allikevel, at lydstyrke er en meget dominant faktor.

Ved posisjoner borte fra sweet spot kan man se for seg at det kan oppstå en endring av dominant indikator som en konsekvens av maskering/filtrering av andre indikatorer. Teoretisk kan man se for seg dette skje selv i posisjoner i kort avstand fra sweet spot, men man må anta at det primært vil skje i posisjoner et godt stykke borte fra høyttalerne. Størst mulighet er det for at det skjer i rom som er fysisk atskilt fra høyttalerne hvor lyden har vært obstruert av fysiske objekter.

Lyd som hindres av fysiske objekter, for eksempel dører, vil bære preg av kraftig filtrering. Høyere frekvenser filtreres generelt kraftigere enn lavere frekvenser. En konsekvens blir at lydkilder med mest energi i lavere frekvensområder vil overleve bedre enn de med mest energi i de høyre områdene. En slik filtrering vil generelt ramme tilført romklang ekstra sterkt. Tidlige refleksjoner som simulerer dybdeplassing i sweet spot kan være borte idet lyden treffer lytteren. Man kan i så måte se for seg at dette endrer innbyrdes posisjon mellom enkelte lydkilder.

Det er muligens noe uinteressant å se på overlevelsen til spatiale trekk der lyden har vært gjenstand for så kraftig filtrering, men det vil uansett være en situasjon som er høyst relevant i konsumenters omgang med musikk.

Det er vanskelig å kvantifisere opplevd avstand i reprodusert lyd. I breddedimensjonen er det en rekke aspekter som kan kvantifiseres med utgangspunkt i avstanden mellom høyttalerne og størrelsen på lytterrommet. Slik er det ikke i avstandsbedømmingen. Høyttalere plassert ved en yttervegg kan enkelt formidle lydkilder som i avstand synes å være plassert hos naboen over gata. Avstandsbedømmingen i reprodusert lyd dreier seg hovedsakelig om innbyrdes relasjoner mellom lydkildene i lydbildet, samt basale opplevelser som ”nært” og ”fjernt”.

Hvilke avstandstrekk som overlever ved posisjoner utenfor sweet spot og hvilke som ikke gjør det, kan koples til de fysiske attributtene hørselen benytter i sin analyse av lyden. I

hvilken grad de fysiske attributtene beholder sin karakter, avgjør også i hovedsak om de spatiale trekkene overlever. Foregående setning er for så vidt gyldig for alle spatiale trekk, men er spesielt tydelig i avstandsbedømmingen.

Hvis vi nå ser på de tre aspektene som formidler avstand – lydstyrke, høyfrekvent innhold og romklang – og hvordan avstandsrelasjonene mellom lydkildene opprettholdes i posisjoner utenfor sweet spot, vil det være romklangen som er gjenstand for størst grad av usikkerhet.

Årsaken er at endringer med hensyn til lydstyrke og høyfrekvent innhold som følge av større reiselengde, vil påvirke de ulike lydkildene likt. Man kan se for seg at tidlige refleksjoner fra lytterrommet forsterker lydstyrken til enkelte lydkilder i større grad enn andre som følge av varierende absorpsjonsgrad i ulike frekvensområder, men da beveger man seg ned på et mikronivå. Endringer i opplevd avstand som følge av stående bølger er kanskje mer relevant, men det vil høre til sjeldenhetene og vil i det minste forutsette relativt statiske lydkilder.

Romklang som avstandsindikator er som tidligere beskrevet fokus for lyttforsøket som ble utført i forbindelse med denne oppgaven. Temaet vil bli ytterligere beskrevet senere, men vi skal allikevel ta et mer generelt blick her.

Den innspilte romklangen vil potensielt være gjenstand for maskering som følge av refleksjoner fra lytterrommet. Det gjelder i særdeleshet tidlige refleksjoner i lydbildet. De sene refleksjonene vil være noe mer skånet for lytterrommets påvirkning, men det avhenger naturligvis av karakteristikken til romklangen i både innspilling og lytterrom.

I kapittel 2.3.1 så vi at laterale refleksjoner bedre formidler avstand enn frontale refleksjoner. I posisjoner utenfor sweet spot vil det generelt oppstå en forringelse/fordreining av refleksjonenes karakteristik. Man må derfor generelt kunne anta en forringelse/innsnevring i opplevd avstand til lydkilder som er tilført laterale tidlige refleksjoner.

#### **4.4 Typiske trekk ved konsumenters lytting**

Det er som tidligere nevnt naturlig å bruke den optimale lyttesituasjonen som målestokk for de endringer som oppstår utenfor denne situasjonen. Tar man med alle faktorer som er med på å definere en virkelig optimal lyttesituasjon er den definitivt ikke lett å oppspore. Den synes å være et evig mål for både musikkprodusenter og hifi-entusiaster. Det hersker relativt stor uenighet rundt flere aspekter, som betydningen av ulike komponenter og deres utforming, men det synes å være relativt stor enighet om at man vanskelig kan oppnå en optimal



lyttesituasjon i det jevne konsumentrom. Hvis det er vanskelig for den skolerte å oppnå noe tilnærmet det optimale i sin egen stue, sier det seg selv at den jevne konsument sjelden er i nærheten.

Konsumentens lyttesituasjon må generelt antas å være preget av tilfeldighet med hensyn til lytterens posisjon i forhold til høyttalerne, og ikke minst høyttalernes plassering i seg selv. Man må anta at praktiske og estetiske valg med hensyn til innredning går streben etter gode lytteforhold en høy gang i de fleste hjem. Avviket fra det optimale manifesterer seg på en rekke områder, både i og utenfor hjemmet. Vi har allerede sett på noen aspekter som plasserer seg på veien fra innspillingsmedium, via reproduksjonsanlegg og fram til lytteren. I tillegg er det en rekke utenforliggende faktorer som potensielt vil påvirke lytterens opplevelse av spatiale trekk. Støy er antakeligvis den faktoren som åpenbarer seg tydeligst. I de resterende avsnittene i dette kapittelet vil jeg trekke fram tre separate aspekter man må anta påvirker opplevelsen av spatiale trekk i høy grad. De tre trekkene er høyttalerplassering, begrenset båndbredde og avspillingsvolum.

#### **4.4.1 Høyttalerplassering**

Betydningen av høyttalernes posisjon i lytterrommet ble så vidt nevnt i avsnittet om breddedimensjonen. Det optimale forutsetter en meget presis symmetri mellom høyttalere og lytter, men det stanser ikke der. Det optimale forutsetter også en symmetri mellom høyttalere og lytterrom. Den presise symmetrien mellom høyttalere og lytter i det optimale forrykkes lett ved at lytteren endrer posisjon. Symmetrien mellom høyttalere og lytterrom forrykkes åpenbart ved usymmetrisk plassering av høyttalerne i forhold til omgivelsene. Symmetrien forrykkes også ved usymmetrisk innredning av selve lytterrommet. I konsumentrom kan man anta at normen er en kombinasjon av begge de to sistnevnte.

Hensynet til øvrig innredning synes å prege høyttalerplasseringen i mange hjem. Det finnes lite statistikk på dette området, så den generaliseringen som skjer er basert på egne observasjoner, samt observasjon av andres observasjoner av samme problemstilling. Det er ikke uvanlig å komme over hjem med en stor usymmetri mellom høyttalere og lytterrom. Stuer der den ene høyttaleren for eksempel står i en bokhylle, mens den andre står plassert i et hjørne bak sofaen. Inntoget av surround-anlegg i konsumenters hjem har antakeligvis medført en viss grad av forbedring på dette området, men mangfoldet er stort.

Usymmetrien i fysisk utforming og plassering vil innebære en usymmetri i de lydsignalene som oppstår i interaksjonen mellom den enkelte høyttaler og rommet. I hvilken

grad spatiale trekk overlever, avgjøres fremdeles av hvorvidt de fysiske attributtene som formidler det spatiale trekket har overlevd når det treffer lytteren.

Forutsatt at høyttalerne spiller i lik fase, altså at de peker i samme retning (det er ikke nødvendigvis gitt), er sannsynligheten stor for at man vil oppleve ulik lydstyrke fra de to høyttalerne. Ulikhetene i lydstyrke vil være en følge av ulik reiselengde for lydbølgene mellom høyttalere og lytter, men også karakteristikken til de tidlige refleksjonene som genereres ut fra signalet til den enkelte høyttaler.

I likhet med mye annet i dette feltet behandler litteraturen/forskningen i liten grad problemet utover situasjoner som er relativt tett knyttet til optimal høyttalerplassering. Den generelle beskrivelsen er også relativt unyansert og begrenser seg til for liten eller for stor vinkel mellom lytter og høyttalere. Ved for liten vinkel beskrives lytteopplevelsen å tendere mot opplevelsen av mono reproduksjon. Mens en for stor vinkel mellom lytter og høyttalere, beskrives med at lytteren hører to distinkte lydkilder, og at det oppstår et "hull i midten" i lydbildet (Rossing, Moore, Wheeler, 2002, s. 578). Det er for så vidt ikke noe galt i selve beskrivelsen, annet enn at den er meget overflatisk.

#### **4.4.2 Begrenset båndbredde**

Begrenset båndbredde er et annet utpreget trekk ved en rekke konsumentanlegg. Vi lytter til musikk fra høyttalere som er både små, store, mellomstore, bærbare og innebygd i bærbare datamaskiner, for å nevne noe. Den begrensede båndbredden som preger konsumentanlegg går først og fremst utover reproduksjonen av lavere frekvenser. I tillegg berører det høyere frekvenser mer enn mellomtone. Små høyttalerelementer er generelt ikke i stand til å formidle lave frekvenser. Mangelen på lave frekvenser innebærer størst forringelse av lydkilder som inneholder mye lave frekvenser. Man må anta at formidlingen av spatiale trekk generelt vil lide under begrenset båndbredde. Men i hvilken grad, avhenger veldig av den øvrige karakteristikken til den enkelte lydkilde.

#### **4.4.3 Avspillingsvolum**

Opplevelsen av spatiale trekk avhenger i stor grad av hvilken lydstyrke musikken reproduseres i. Avspillingsvolumets betydning gjelder naturlig nok også den optimale lyttesituasjonen, men det nevnes i denne delen da man av erfaring vet at konsumenters musikklytting dekker hele desibelskalaen. En grunnleggende og åpenbar forutsetning for å kunne oppleve et spatiale trekk i reprodusert lyd er at lyden som danner trekket må være hørbar. Lytting til musikk med lavt avspillingsvolum vil derfor generelt avdekke færre spatiale trekk. I den andre enden av skalaen ligger for høye avspillingsvolum. Et hvert

lytterom har sitt metningspunkt med hensyn til lydstyrke. Overstiger den samlede lydstyrken dette metningspunktet vil spatiale trekk generelt forringes. Ved høye avspillingsvolum vil lyden også være mer utsatt for forvrengning i reproduksjonsanlegget.

## 5 Beskrivelse av lytteforsøk

I dette kapittelet skal jeg beskrive arbeidet med å forberede og gjennomføre det praktiske lytteforsøket som er en del av grunnlaget for denne oppgaven. I deler av dette kapittelet benyttes en rekke lydtekniske begreper. Sentrale begreper er beskrevet i begrepsavklaringen i kapittel 1.7, men det benyttes enkelte ord som ikke defineres. Ordene knytter seg til omtalen av lydteknisk utstyr som er anvendt og tekniske grep utført i miksesituasjonen. Jeg antar at lesere som er interessert i utstyr og tekniske justeringer, allerede har kjennskap til denne type termer. Det er en del av tradisjonen å ta med detaljer av denne art, men detaljene kan overses uten å gå glipp av noe vesentlig.

Sammen med hovedoppgaven er det vedlagt en cd (se Appendiks B). På denne er det lagt inn en kopi av lytteforsøkets oppgaver slik de ble presentert for deltakerne.

### 5.1 Innspilling i Rakkestad

For å ha full kontroll og oversikt over den romklangen som er til stede i innspilt musikk er det nødvendig å ta utgangspunkt i materiale som er ekkofritt, altså helt uten romklang. Noen former for ekkofritt lydmateriale kan genereres syntetisk. Man kan også fange inn ekkofrie signaler fra elektriske instrumenter ved å plugge dem rett i en miksepult, men for å fange inn lyden fra akustiske instrumenter, og ikke minst menneskets stemme, er man nødt til å gjøre opptak i såkalt ekkofrie rom.

Det finnes en mengde ekkofrie innspillinger tilgjengelig på nettet, både gratis og til salgs, men det er en slående mangel på flerspors materiale som til sammen utgjør en musikalsk helhet. Videre synes hovedandelen av ekkofrie opptak å bestå av klassisk musikk. Fokus for denne oppgaven er opplevelse av populærmusikk. Det er derfor også en

forutsetning at lydeksemlene i forsøket faller inn under denne kategorien. I forberedelsene til lytteforsøket var det derfor nødvendig å få tilgang til et ekkofritt rom hvor jeg kunne spille inn råmateriale til lydeksemlene.

Det er langt mellom ekkofrie rom i nærheten av hovedstaden, og generelt vanskelig for en hovedfagstudent i musikk å få tilgang til de få som er. Jeg var allikevel så heldig å komme i kontakt med en hyggelig privatperson i Rakkestad som har bygd et slikt rom hjemme hos seg selv. Han lot meg velvillig disponere rommet en hel dag. Med samboer som maskinoperatør, kvinnelig sanghjelp fra Fredrikstad og undertegnede som mannlig vokalist og instrumentalist, fikk jeg spilt inn råmaterialet jeg trengte til forsøkene.

### **5.1.1 Det ekkofrie rommet**

Det ekkofrie rommet i Rakkestad er egentlig bygd for testing av hifi-utstyr og er derfor relativt lite av størrelse. Det er allikevel mer enn stort nok til opptak av gitar og vokal ved nærmikking, slik det ble gjort til dette forsøket.

Rommet har akseptabel demping helt ned til 100 Hz, som er meget bra for et rom av denne størrelsen. Det avviker noe fra industriell standard med henblikk på lydlekkasje fra utsiden av rommet, men dette utgjorde ikke noe problem under innspillingen.

### **5.1.2 Innspillingsutstyr**

Opptakene ble gjort med følgende utstyr:

- Mikrofoner: 1 stk AKG C 414 B-ULS  
1 stk Neumann KM100, kapsel (40) nyrekarakteristikk
- Lydkort: M-Audio FireWire 410
- Datamaskin: Acer Aspire 5000 bærbar
- Monitoring: Genelec 8020 A
- Programvare: Cubase SX

Det ble gjort opptak av to akustiske gitarer (nylon og stålstrengs) og både dame- og herrestemme. Begge mikrofonene ble benyttet ved alle opptakene. Ved vokalopptakene ble AKG-mikrofonen stilt inn på omni-karakteristikk. Opptak av gitarene ble det gjort ved såkalt ms-teknikk, hvor AKG-mikrofonen var stilt inn med 8-talls-karakteristikk. Ved miksing til forsøket ble det allikevel bare benyttet signal fra en enkelt mikrofon fra hver lydkilde. Mer om dette nedenfor.

### 5.1.3 Musikere og instrumenter

Både den musikalske- og lydtekniske kvaliteten på populærmusikkinnspillinger avgjøres i stor grad av musikernes framføring av musikken. Det faktum ble ikke mindre tydelig ved innspilling i ekkofrie omgivelser i Rakkestad. Undertegnedes gitarferdigheter, samt et snev av tidspress, innebar at innspillingen ble langt fra tilfredsstillende vurdert ut fra undertegnedes subjektive preferanser. Kvinnevokalen var derimot upåklagelig. Til tross for det noe stramme og lite dynamiske uttrykket ble innspillingen vurdert å være brukende til lytteforsøket.

Det ble spilt inn tre ulike musikkstykker:

1. She's got her ticket (Tracy Chapman): A cappella kvinnestemme
2. Blackbird (Lennon/McCartney): Herrestemme og stålstrengs gitar
3. Here There and Everywhere (Lennon/McCartney): Kvinnestemme, nylonstrengs gitar og elektrisk bass. Den elektriske bassen ble spilt inn i forkant av sesjonen i Rakkestad.

## 5.2 Miksing

Etter innspilling ble materialet bearbeidet videre i ProTools. Arbeidet foregikk i studio Wiggen hos NOTAM (Norsk nettverk for Teknologi, Akustikk og Musikk). Utstyret som ble benyttet i studio Wiggen var:

- ProTools HD 3 accel system
- Dynaudio Acoustic Air series stereo lytting
- Sony Oxford EQ + filters
- Waves Rcompressor (kompresjon)
- Digidesign DynIII De-esser (frekvensspesifikk kompresjon)
- Waves TrueVerb (romklangsimulering)

Det ble ikke utført noen form for panorering under miksing, mens eq og kompressorer ble lagt inn som inserts. Kompressor lå før eq i signalgangen, der begge ble benyttet. Tabell 5-1 viser en oversikt over de endringer som ble gjort under miksing.

Innspilling/instrument (mikrofon i parentes)	Effekt	Parameterendring
<b>She's got her ticket</b>		
Vokal (AKG C 414 B-ULS)	EQ	MF = 1499.1 Hz, Q = 1.99, +1.79 dB HF = 6317.4 Hz, Q = 0%, +5.74 dB
<b>Blackbird</b>		
Vokal (Neumann KM100)	EQ	Lowcut = 50.1 Hz 12dB per octave HF = 6671.3 Hz, Q = 0 %, +8.97 dB
Gitar (Neumann KM100)	EQ	LF = 122.9 Hz, Q = 0 %, +4.46 dB HF = 5562.9 Hz, Q = 0 %, +4.77 dB
	De-esser	1.5kHz, -3.0 dB
<b>Here There and Everywhere</b>		
Vokal (AKG C 414 B-ULS)	Kompressor	threshold -9.6, attack 5.02, rel 50.2, ratio 3.35, gain -3.0 (arc, electro, warm)
	EQ	lowcut 70.6 Hz, -12db octave HF = 6610.9 Hz, Q = 0%, +9.21dB
Gitar (Neumann KM100)	Kompressor	threshold -13.1, attack 62.5, release 160, ratio 5.06 (arc, electro, warm)
	EQ	lowcut 55.1 -12db octave, MF = 560.7 Hz, Q = 4.04, -5.59dB HF = 4159.6 Hz, Q = 0%, +5.08dB
Bass (DI)	Kompressor	threshold -9.8, attack 16.0 ms, release 160 ms, ratio 3.7ms (arc, electro, warm)
	EQ	MF = 1141.4Hz, Q = 2.51, +4.63dB

**Tabell 5-1:** Oversikt over lydtekniske justeringer som ble utført under nedmiksing av lydeksempelene til lytteforsøket. Parameterendringene refererer til navnene på parameterne slik de er benevnt i innstikkene (plugin) som er benyttet. Følgende forkortelser benyttet: HF = High Frequency, MF = MidFrequency, Q = båndbredde, DI = direct insert (= plagget direkte inn i miksepulten)

### 5.2.1 Lydeksempelene

Av de tre innspilte sangene ble det lagd fem ulike lydeksempler. Hvert lydeksempel ble deretter mikset i fire ulike versjoner. Forskjellen mellom de fire versjonene ble skapt ved å variere lydstyrken på de tidlige refleksjonene som ble tilført vokalen.

Samlet utgjorde de fire ulike versjonene av hvert lydeksempel en enkelt oppgave i forsøket. Med andre ord besto forsøket av fem ulike oppgaver. Hver av de fem oppgavene var basert på ett av de fem lydeksempelene, og inneholdt fire ulike versjoner som deltakeren skulle vurdere opplevelsen av.

I utvelgingen av lydeksempelene ble det lagt vekt på at disse skulle representere en helhetlig musikalsk frase. Det ble også lagt vekt på at eksempel materialet skulle være variert med hensyn til både transientrikdom og kompleksitet i lydbildet. Vår evne til å oppleve tidlige refleksjoner er generelt høyere når refleksjonene følger lydkilder som er transientrike og stadig skiftende. Høyere kompleksitet i lydbildet vil generelt forringe evnen til å oppleve refleksjonene, og vil dermed også utfordre opplevelsen av spatiale trekk.

Hvert av lydeksempelene hadde en varighet på omtrent 12 sekunder. Tabell 5-2 viser en oversikt over de fem lydeksempelene/oppgavene. Kolonnen med overskriften "algoritme",

refererer til algoritmen i den kunstige romklengen som ble tilført vokalen. Romklengen blir behandlet i kapittel 5.2.2.

Oppgave	Lydeksempel	Algoritme	Instrumenter	Versjon 1	Versjon 2	Versjon 3	Versjon 4
1	She's got her ticket utdrag 1	mhII	Vokal	+3 dB	0 dB	-6 dB	-12 dB
2	Blackbird utdrag 1	stA	Vokal	+6dB	-3dB	-9 dB	Tørr
3	Blackbird utdrag 2	stA	Vokal og gitar	+3 dB	0 dB	-6 dB	-12 dB
4	She's got her ticket utdrag 2	mhII	Vokal	+6dB	-3dB	-9 dB	Tørr
5	Here There and everywhere	mhII	Vokal, gitar og bass	+3 dB	0 dB	-6dB	-12dB

**Tabell 5-2:** Oversikt over oppgavene i forsøket – med tilhørende lydeksempler og nivåer på de tidlige refleksjonene i de ulike versjonene. Desibeltallene refererer til nivået på de tidlige refleksjonene. Referanseverdien 0 dB er hentet fra nivåforholdet mellom direkte lyd og tidlige refleksjoner slik det er i klangealgoritmens grunnoppsett. "Tørr" betyr ingen tilførte tidlige refleksjoner.

#### *Oppgave 1: She's got her ticket – utdrag 1*

Oppgave 1 består av en kvinnestemme som synger Tracy Chapman-låten "She's got her ticket". Innspillingen er også gjort a cappella, uten noen form for akkompagnement av musikeren. Utdraget som benyttes i oppgaven bærer preg av å være rikt på transienter og luft. Det dynamiske preget, i kombinasjon med et lite komplekst lydbilde, må antas å formidle variasjoner i avstand meget bra.

De tidlige refleksjonene er tilført med nivåene +3 dB, 0 dB, -6 dB og -12 dB. Referanseverdien 0 dB tar som nevnt utgangspunkt i nivåforholdet mellom direkte lyd og tidlige refleksjoner, slik det er i romklangprogrammets grunninnstillinger.

#### *Oppgave 2: Blackbird – utdrag 1*

Oppgave 2 består av en herrestemme som synger refrenget på Beatles-låten "Blackbird". Sangstemmen ble opprinnelig spilt inn akkompagnert av gitar. Utdraget er valgt ut fordi det står i kontrast til oppgave 1. Sammenlignet med oppgave 1, preges oppgave 2 av relativt få transienter, utdratte toner og lite luft. Man kan anta at lydbildet således formidler variasjoner i avstand i mindre grad enn oppgave 1.

De tidlige refleksjonene er tilført med nivåene +6 dB, -3 dB, -9 dB og tørr. Versjonen som er tørr inneholder ingen tidlige refleksjoner. Det er generelt større gap mellom de ulike nivåene her enn i oppgave 1. Versjonen hvor de tidlige refleksjonene er forsterket +6 dB må sies å være unaturlig høyt.



### *Oppgave 3: Blackbird– utdrag 2*

Oppgave 3 er hentet fra samme innspilling som oppgave 2, men her akkompagneres herrestemmen av stålstrengs gitar. Det er kun vokalen som er tilført romklang. Det er ikke foretatt noen form for panorering. Utdraget er hentet fra starten av verset. Karakteren til stemmen er langt rikere på transienter og luft. Samtidig øker kompleksiteten i lydbildet ved innførselen av gitaren. Den er antatt at den økte kompleksiteten i lydbildet kompliserer avstandsbedømmingen for deltakerne i forsøket.

De tidlige refleksjonene er tilført med nivåene +3 dB, 0 dB, -6 dB og -12 dB. Dette er de samme nivåene som i oppgave 1.

### *Oppgave 4: She's got her ticket – utdrag 2*

Oppgave 4 er hentet fra samme innspilling som oppgave 1, og er fremdeles a cappella kvinnestemme. Sammenlignet med oppgave 1, bærer eksempelet preg av litt færre transienter og litt mer utdratte toner, og dermed litt mindre luft.

De tidlige refleksjonene er tilført med nivåene +6 dB, -3 dB, -9 dB og tørr. Nivåene er de samme som i oppgave 2, og har dermed litt større gap mellom de ulike nivåene, sammenlignet med oppgave 1. Man kan anta at et større gap mellom nivåene bedre formidler forskjeller i avstand.

### *Oppgave 5: Here there and everywhere*

Musikken i oppgave 5 er en innspilling av Beatles-låten "Here there and everywhere" og representerer i høy grad det mest komplekse lydbildet. I tillegg til nylonstrengs gitar, akkompagneres vokalen i tillegg av el-bass. Det er kun vokalen som er tilført romklang. Gitaren spiller fingerspill med underdeling på åttendedeler, mens bassen er noe mer utdratt. De hyppige anslagene til gitaren, bidrar til et mer komplekst lydbilde. Det samme gjør naturligvis bassen. Vokalen, som er en kvinnestemme, er for øvrig både rik på transienter og luft.

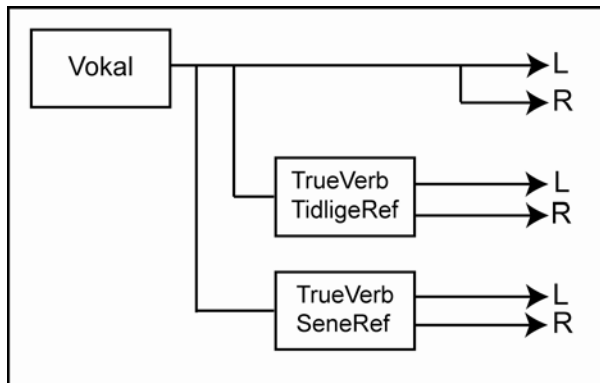
De tidlige refleksjonene er tilført med nivåene +3 dB, 0 dB, -6 dB og -12 dB, og er således likt med oppgave 1 og 3. Det er antatt at dette er den oppgaven som vil skape størst utfordringer for deltakerne i forsøket.

## **5.2.2 Romklengen**

De fire ulike versjonene av hvert lydeksempel er skapt ved å variere styrkeforholdet mellom direkte lyd og kunstige tidlige refleksjoner. Uansett besetning i lydeksempelene er det kun vokalen som tilføres romklang. Det er utelukkende styrkeforholdet mellom direkte lyd og

tidlige refleksjoner som endres mellom de ulike versjonene. Den samme vokalen blir også tilført sene refleksjoner, men disse forblir helt like mellom versjonene.

Ved standard mikseoppsett for tilføring av romklang vil en endring av tidlige refleksjoner også innebære en endring av de sene refleksjonene. For å unngå at så skjedde, ble det under nedmiksing benyttet en signalgang som avviker noe fra det vanlige. Figur 5-1 viser et skjema over signalgangen.



**Figur 5-1:** Signalgang ved tilføring av romklang til vokal under forberedelsene til lytteforsøket

I de ferdig miksedde lydeksempelene er det tatt utgangspunkt i to ulike romklangalgoritmer i TrueVerb. TrueVerb er et innstikk (*engelsk: plugin*) som simulerer romklang i ProTools. De to romklangalgoritmene som ble benyttet var Studio A (forkortet: stA) og Medium Concert II (forkortet: mhII). Begge algoritmene har i utgangspunktet en relativt kort etterklangstid (RT60), hvor stA = 0,8 sek og mhII = 1,3 sek. I hovedsak ble grunninnstillingene i programmet benyttet, med unntak av parameterne distance (avstand) og predelay (initial forsinkelse). Romklangalgoritmene ble valgt på bakgrunn av undertegnedes egne estetiske vurderinger.

#### *Avstandsparameteren (distance)*

Denne parameteren avgjør den simulerte avstanden mellom lydkilde og lytter i algoritmen. Endring av verdien på parameteren innebærer en omvendt proporsjonal påvirkning av styrkeforholdet mellom direkte lyd og romklang. Med andre ord, jo lenger simulert avstand desto kraftigere romklang. I TrueVerb innebærer en oppjustering av avstandsparameteren til et visst punkt også en flatere omhyllingskurve mellom første og siste tidlige refleksjon. I tråd med forsøkets intensjon ble avstandsparameteren justert opp for å ha et utgangspunkt som simulerer mest mulig avstand.

### *Initial forsinkelse (predelay)*

Denne parameteren styrer forsinkelsen mellom starten på den direkte lyden og starten på de sene refleksjonene. Parameteren ble justert opp for å unngå en overlapping mellom de tidlige- og sene refleksjonene. I praksis innebar det en forsinkelse på 80 ms for stA og 100 ms for mhII. I begrepsavklaringen i kapittel 1.7 så vi at ”predelay” helst benyttes i omtalen av den initiale forsinkelsen mellom direkte lyd og første tidlige refleksjon. Det er ikke den parameteren som omtales her. Parameteren som omtales her styrer forsinkelsen på de sene refleksjonene.

### *Nivåforhold direkte lyd og tidlige refleksjoner*

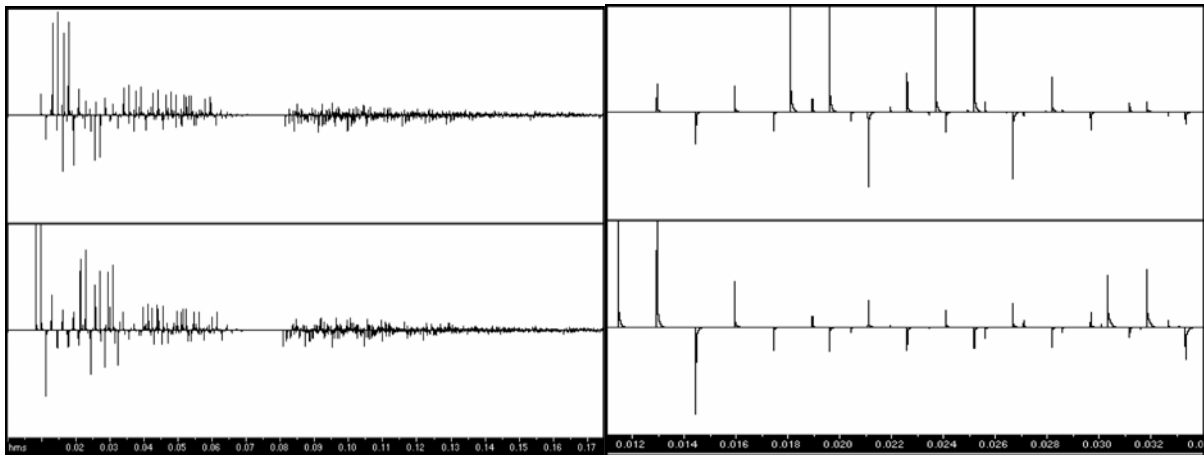
Jeg skal gå nærmere inn på forskjellen mellom de ulike versjonene av lydeksemplene lenger nede, men de er som nevnt generert med ulikt nivå på de tidlige refleksjonene som er tilført vokalen. Utgangspunktet for nivåjusteringen mellom versjonene ble gjort ut fra impulsresponsen til de romklangalgoritmene jeg benyttet. Ved å finne nivåforholdet mellom direkte lyd og første refleksjon i impulsresponsen, kunne jeg justere signalfadere i ProTools deretter.

### *De sene refleksjonene*

De sene refleksjonene ble holdt på samme nivå i alle versjonene, uavhengig styrken på de tidlige refleksjonene. For å fremstå for dominante i lydbildet, ble nivået på de sene refleksjonene senket med -9 dB i forhold til algoritmens grunninnstilling.

## **5.2.3 Impulsrespons**

Figur 5-2 viser impulsresponsen til romklangalgoritmen Studio A, med et styrkeforhold mellom de tidlige- og sene refleksjonene slik det var i utgangspunktet. Nivåforholdet mellom refleksjonene og den direkte lyden kommer ikke tydelig fram i impulsresponsen. Den kraftigste tidlige refleksjonen, som var andre refleksjon i høyre kanal, lå -14 dB under den direkte lyden. Den observante leser vil se at det er konstruert et tydelig skille mellom de sene og tidlige refleksjonene. Dette er noe uvanlig i naturen, men er slettes ikke uvanlig innen populærmusikkproduksjon. Man kan også se ut fra impulsresponsen at det er lav grad av korrelasjon (samsvar) mellom refleksjonene i de to kanalene. Den lave graden av korrelasjon kommer enda tydeligere fram i figur 5-3 som er et utsnitt av de første refleksjonene i impulsresponsen til Medium Concert II. Den lave korrelasjonen bidrar til å simulere laterale refleksjoner, som vi vet bidrar i positiv retning til en rekke spatiale trekk, deriblant avstand. Impulsresponsene viser for øvrig at ankomsttidspunktet for første refleksjon var 8 ms i Studio A og 11 ms i Medium Concert II.



**Figur 5-2:** Impulseresponsen til romklangalgoritmen Studio A

**Figur 5-3:** Utsnitt av impulsresponsen til Medium Concert II – De første tidlige refleksjonene

### 5.2.4 Lydstyrken på de tidlige refleksjonene

Lydstyrken til de tidlige refleksjonene kan synes å være noe tilfeldig valgt. Det er de for så vidt også. Det kunne vært interessant å basere lydstyrken på avdekkede grenseverdier, men ble ikke gjort av to årsaker. For det første er det vanskelig å ta tak i de grenseverdiene som er påvist, da disse forutsetter meget kontrollerte omgivelser. Man kunne muligens ha basert nivåene på slike grenseverdier og tilført gradvis mer, men man ville da ikke fjernet inntrykket av tilfeldig valgte verdier uten å ha gjennomført grundige laboratorieforsøk på forhånd. Den andre årsaken er at utgangspunktet for denne oppgaven er musikkproduksjon, og innenfor det feltet må man kunne anta at det sjelden tilføres kunstige tidlige refleksjoner helt ned mot grenseverdiene. Nivåforskjellene mellom de ulike versjonene ble i stedet valgt med utgangspunkt i steg på 3dB, som gir enkle forholdstall i lydnivå.

### 5.2.5 Det som ikke ble gjort

Den eneste forskjellen mellom versjonene var som nevnt en endring av nivåforholdet mellom direkte lyd og tidlige refleksjoner. I naturlige omgivelser vil en økning i avstanden innebære at den direkte lyden i praksis blir høypassfiltrert. Av åpenbare årsaker er ikke forsøksmaterialet filtrert på denne måten. Selv om utelatelsen av høypassfiltrering er i strid med naturen, må det anses å være i tråd med populærmusikkproduksjonen. Det hører nok til sjeldenhetene i populærmusikken at hovedvokalen blir høypassfiltrert.

### 5.2.6 Nivåjustering

En av de store utfordringene i forberedelsene til et forsøk av denne art, knytter seg til justeringen av avspillingsnivået til musikk eksempelen som benyttes i forsøket. Den opplevde lydstyrken til et lydsignal (*engelsk: loudness*) kan ha stor betydning for vår opplevelse av det

(Rumsey, Mason, 2002, s.7). Dette faktum blir særdeles tydelig når man arbeider med lydsignaler som er tilført tidlige refleksjoner med ulik lydstyrke.

I lydproduksjon bruker man gjerne benevnelsen "tørt" om ubehandlede lydsignaler. Det motsatte er et "vått" signal. I denne sammenheng, lydsignaler som ikke er tilført kunstig romklang. Hvis vi tar utgangspunkt i det tørre signalet og gradvis tilfører mer og mer tidlige refleksjoner, vil vi gradvis snevre inn nivåforholdet (ratioen) mellom direkte lyd og tidlige refleksjoner. En innsnevring av nivåforholdet innebærer en økning i opplevd avstand til lydkilden, forutsatt at andre faktorer forblir like. Men alle andre faktorer forblir ikke like i det man beveger volumkontrollen til de tidlige refleksjonene oppover.

Etter hvert som man tilfører mer og mer tidlige refleksjoner, øker man samtidig den opplevde lydstyrken markant. Den initiale lydstyrken på de ulike versjonene av hvert lydeksempel var så markant at da det ujusterte materialet ble testet for opplevd endring i avstand på to personer, var vurderingen til personene diametralt forskjellig fra hva forholdstallene mellom direkte lyd og tidlige refleksjoner tilsa.

Det finnes per i dag ingen anerkjent metode for måling av opplevd lydstyrke på dynamisk lydmateriale. En kombinasjon av menneskelig vurdering og Moores (m.fl.) modell for statisk materiale (Moore, Glasberg, Baer, 1997), synes å fungere bra, også på dynamisk materiale (Rumsey, Mason, 2002, s. 7). Da undertegnede ikke hadde tilgang til gode verktøy til å benytte denne metoden, ble justeringen kun utført på bakgrunn av menneskelig vurdering. En slik tilnæringsmåte er fremdeles anerkjent som en gyldig metode (Rumsey, 2006).

For å holde den opplevde lydstyrken på et jevnt nivå, må nødvendigvis noe justeres ned når de tidlige refleksjonene justeres opp. Det følger naturlig av sammenhengen at den skadelidende i dette lyd materialet må være den direkte lyden fra vokalen. I praksis ble nivåjusteringen utført ved å først lage en miks av vokal og tidlige refleksjoner. Nivået på vokalmiksene ble deretter jevnet ut ved å normalisere dem digitalt. I normaliseringsalgoritmen ble det benyttet en såkalt 'equal loudness contour', som er en litt mer frekvensanalyserende normaliseringsmetode som har til hensikt å flate ut ulikheter i opplevd lydstyrke. Etter normaliseringen ble vokalmiksene mikset med sene refleksjoner og andre instrumenter der det var det. Til slutt ble nivåene på miksene ytterligere justert ved undertegnedes egen lytting i forsøkslokalet.

### **5.2.7 Feil på nivåer i grensesnittet**

Til tross for den møysommelige prosessen med nivåjustering oppsto det på veien mellom lydredigeringsprogram og grensesnittet som forsøkspersonene benyttet en endring av lydnivået på to versjoner. Årsaken til feilen skyldes ene og alene slurv av undertegnede.

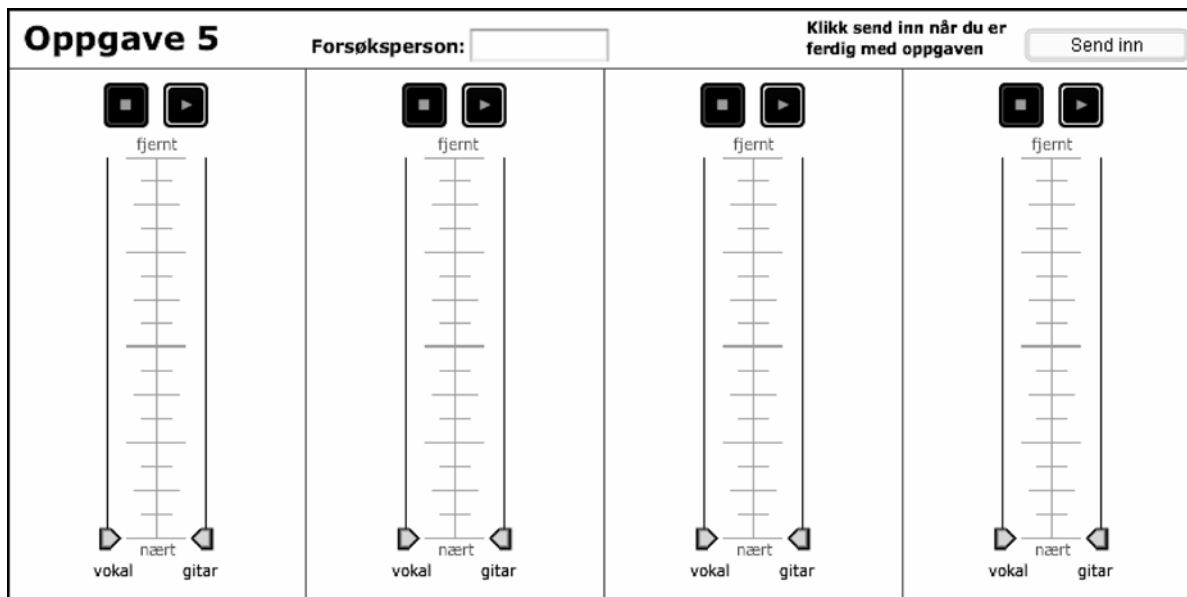
Feilen ble ikke oppdaget før tredje forsøksperson hadde gjennomført lytteforsøket. Konsekvensene av feilen ble vurdert til å ikke ha noen fundamental betydning for resultatet av forsøket som helhet, og det ble derfor besluttet å fortsette med samme lydmateriale på resten av deltakerne. Som vi skal se litt senere, er konsekvensen av nivåfeilene meget påtagelig, og kunne kanskje i kunnskapsformidlingens navn vært presentert som en tilsiktet illustrasjon av den betydningen opplevd lydstyrke har for avstandsvurderingen. De to versjonene er versjon 4 i oppgave 2, og versjon 3 i oppgave 4. De to versjonene er følgelig ikke tatt med i grunnlaget for analysen.

## **5.3 Grensesnittet**

### **5.3.1 Flash-applikasjon**

Til gjennomføringen av lytteforsøket ble det lagd et grafisk grensesnitt som forsøkspersonene benyttet i besvarelsen av oppgavene. Grensesnittet er utviklet i Flash, med lagring av dataene i en MySQL-database via PHP. Teknologien er valgt fordi det er en lett tilgjengelig teknologi, som både er nettbasert og rask å utvikle i. At det er nettbasert er en fordel med tanke på mangfoldiggjøring av hovedoppgaven. Håpet er å få publisert en publikumstilpasset versjon med ”svar” og forklaringer i forlengelsen av dette arbeidet. I tillegg vil også kildekoden til applikasjonen publiseres, til fri benyttelse av andre.

Figur 5-4 viser det grafiske grensesnittet som ble presentert for deltakerne. Grensesnittet til hver enkelt oppgave var, som man ser, delt inn i fem deler. Den øverste delen av grensesnittet består av overskrift, felt for utfylling av navn og knapp til innsending av oppgaven. I tillegg er det fire felt med knapper for avspilling og vurdering/sammenligning av de fire versjonene av hvert lydeksempel. I oppgaver hvor det kun er vokal som skal vurderes, er det også bare skyveknapper for vokal.



Figur 5-4: Grensesnittet som ble benyttet i forsøket

### 5.3.2 Graderingsskala

Deltakernes besvarelse blir formidlet i grensesnittet ved å bevege skyveknappene for hvert instrument oppover eller nedover på graderingsskalaen man ser på figur 5-4. Skalaen er 256 piksler høy, hvor posisjonen til skyveknappen også blir lagret med denne oppløsningen.

Mellom skyveknappene er det lagt inn hjelpelinjer med en avstand på 16 piksler mellom hver linje. Fordi det er vanskelig å kvantifisere avstandsopplevelse i reproduisert musikk, er det ikke benyttet noen form for tallangivelse av skalaen. Ordene ”nært” og ”fjernt” er benyttet for å sikre at besvarelsene ikke blir utført opp ned.

### 5.4 Oppsett i lytterom

Av praktiske hensyn, og i mangel av tilgang til lokaliteter som tilfredsstillende anbefalingene til International Telecommunications Union (ITU-R BS1116), ble lyttforsøket gjennomført i undertegnede egen stue. Rommets utforming, innredning og akustiske karakteristikk må, subjektivt vurdert, kunne sies å være i nærheten av en gjennomsnittlig norsk stue, og representerer i så måte akustiske omgivelser som vil oppleves som ”vanlige” for den jevne nordmann. For ordens skyld må det sies at norske stuer antakeligvis avviker noe fra enkelte andre lands stuer. I England dekker man for eksempel gjerne hele gulvet med teppe. Strengt tatt er resultatene av forsøket Takhøyden på 310 centimeter er kanskje det arkitektoniske attributtet som avviker mest fra et landsdekkende gjennomsnitt, og muligens beveger rommet mer over mot et hovedstadsgjennomsnitt. For øvrig er rommet relativt kvadratisk med sine 490 x 487 centimeter mellom de malte huntonittplatene på veggene. Taket er helsparklet og

malt, mens parketten på gulvet er delvis dekket av et ensfarget teppe fra Ikea. Rommets utforming innebærer noen relativt dominante resonanser, spesielt rundt frekvensene 35, 70, 105 og 141 Hertz. Stuen er innredet med sofa, stuebord, lenestol, 3 bokhyller med bøker, en stereobenk med tv på og en kontrabass i et hjørne. Det er to høye vinduer langs den ene veggen og et spisestuebord langs motsatt vegg, hvor åpningen ut til gangen også befinner seg. Se for øvrig impulsresponsene lenger nede.

Gangen, som ble benyttet i lytteposisjon 3, har form som en speilvendt L. Den er 120 cm bred, 270 cm langs kortenden som åpningen til stuen befinner seg i enden av, og 300 cm lang den andre veien. Det er parkett på gulvet, veggene består av malt mur og fire dører i tre.

#### **5.4.1 Høytalerplassering**

Lydeksempelen ble reprodusert gjennom to Dynaudio BM5A aktive nærfeltsmonitører plassert på stativer 1 m over gulvet. Høytalerne ble plassert langs vindusveggen med 10 centimeters avstand til veggen og 135 cm mellom høytalerne. Venstre høytaler var plassert tett mot midten av veggen, mens høyre høytaler var plassert i forkant av det ene vinduet. Med høytaleroppsettet som ble benyttet, var romradius ca. 180 cm fra høytalernes front. Høytalerplasseringen ble gjort på bakgrunn av to kriterier:

1. Høytalerplasseringen skulle bære preg av en viss grad av tilfeldighet, da den jevne konsument kan synes å velge høytalerplassering ut fra helt andre kriterier enn ønsket om optimale lytteforhold.
2. Det skulle allikevel være mulig å posisjonere lytteren i en tilnærmet optimal posisjon, med noe avstand til alle overflater.

#### **5.4.2 Lytterplassering**

Lytteforsøket ble gjennomført med forsøkspersonene sittende ved tre ulike posisjoner. Disse var sweet spot, ved spisestuebordet og ute i gangen. Ved alle posisjoner hadde forsøkspersonene et bord med en 19" lcd skjerm og trådløs mus foran seg.

##### *Posisjon 1 – sweet spot*

Utgangspunktet for hele forsøket var å lete etter ulikheter i lytteopplevelsen mellom ulike lytteposisjoner. Målestokken for en slik type forsøk bør nødvendigvis være det optimale. Nå oppnår man nok aldri en optimal lyttesituasjon i vanlig innredede stuer, så i praksis blir målestokken det nærmeste man kommer det optimale under de gitte forutsetningene. Posisjon 1 i forsøket ble derfor sweet spot.



Med en avstand på 135 cm mellom høyttalerne, skal sweet spot også ligge 135 cm fra grunnlinjen mellom høyttalerne. Ved kontrolllytting ble posisjonen flyttet noen centimeter nærmere til omtrent 125 cm, da dette gav en bedre stereogjengivelse.

#### *Posisjon 2 – bord*

Et annet utgangspunkt for forsøket var å belyse former for lytteposisjoner man kan anta mange konsumenter befinner seg i når de lytter til musikk. Musikklytting sittende ved et bord må i så måte kunne anses som en utbredt lytteposisjon for konsumenter.

En posisjon ved spisebordet ble funnet egnet for forsøket med bakgrunn i flere faktorer som alle kan anses som vanlige ved konsumenters lytting. Avstand mellom lytteren og nærmeste høyttaler var 400 centimeter.

1. Posisjonen ligger i fjernfeltet og godt til side for akse midt mellom høyttalerne.
2. Det er store overflater tett innpå lytteren. Posisjon er tett inntil bakre vegg og lytteren har samtidig bordet i magehøyde foran seg.
3. Lytteren har det ene øret rettet bort fra høyttalerne og inn mot veggen.

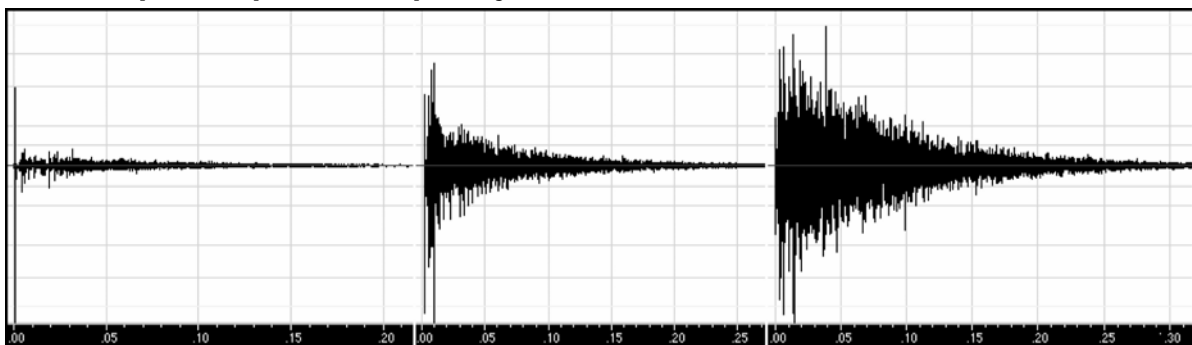
#### *Posisjon 3 – gang*

Posisjon 3 har tilsvarende trekk som posisjon 2, men i tillegg også følgende:

1. Det er ikke fri sikt mellom høyttaler og lytter.
2. Lytteren befinner seg i et annen rom enn høyttalerne.

I posisjon 3 var avstanden mellom lytter og nærmeste høyttaler 650 centimeter.

### **5.4.3 Impulsresponsene av posisjonene**



**Figur 5-5:** Impulsresponsene ved de lytteposisjonene som ble benyttet i forsøket. Til venstre: Posisjon 1 (sweet spot) – I midten: Posisjon 2 (bord) – Høyre: Posisjon 3 (gang)

Når man sammenligner impulsresponsene ved de ulike posisjonene ser man umiddelbart at det er en stor forskjell mellom dem (se figur 5-5). Mens ratioen mellom direkte lyd og refleksjoner er relativt stor ved sweet spot, ser man at dette styrkeforholdet er totalt endret ved

de to andre posisjonene. Man ser til og med at ved et par steder medfører den samlede energien av flere refleksjoner et lydtrykk som er høyere enn den direkte lyden.

#### 5.4.4 Avspillingsnivå

Det generelle avspillingsnivået på lydeksempelene ble valgt med utgangspunkt i det som ble ansett å være et behagelig, men samtidig klart og tydelig lydnivå. Målt ut fra hvit støy i sweet spot, ble nivået justert til 70 dBA. Avspillingsnivået ble justert mellom posisjonene for å opprettholde et jevnt lydtrykk uavhengig av posisjon.

### 5.5 Gjennomføring av forsøket

Den praktiske delen av lytteforsøket ble gjennomført med en deltaker av gangen. Undertegnede var ikke til stede i rommet under selve gjennomføringen, men kom inn for å flytte dataskjerm og mus mellom de tre posisjonene. Avspilling av lydeksempler og deltakernes besvarelse ble i sin helhet utført ved hjelp av dataskjerm og mus i flash-grensesnittet som er beskrevet ovenfor.

#### 5.5.1 Oppgavene

Lytteforsøket besto som tidligere beskrevet av fem ulike oppgaver. Hver oppgave var satt sammen av de fire ulike versjonene av det enkelte lydeksempel. Deltakerne besvarte alle fem oppgavene ved hver av de tre lytteposisjonene – sweet spot, bord og gang. Fem oppgaver ganger tre posisjoner innebar at hver deltaker besvarte 15 oppgaver til sammen.

Det ble lagt vekt på at forsøkspersonene ikke skulle kunne besvare oppgavene ut fra hukommelse av besvarelse fra foregående posisjoner. Plasseringen til de 4 versjonene innenfor den enkelte oppgave ble derfor tilfeldig plassert i grensesnittet, med ny rekkefølge mellom de ulike posisjonene. Det ble samtidig tatt høyde for at rekkefølgen på versjonene ikke skulle gjenspeile den gradvise tilføringen av refleksjoner. Lydeksempelene ble for øvrig presentert for deltakerne i tre ulike rekkefølger, men rekkefølgen på lydeksempelene fra en posisjon til en annen forble lik for den enkelte deltaker. Sitteplassrekkefølgen ble også variert, slik at man fikk en jevn fordeling mellom posisjoner og deltakere, (se tabell 5-3).

Deltaker	1	2	3	4	5	6	7	8
Sitteplass 1	sweet	bord	gang	sweet	gang	bord	sweet	bord
Sitteplass 2	bord	sweet	sweet	gang	bord	gang	bord	gang
Sitteplass 3	gang	gang	bord	bord	sweet	sweet	gang	sweet

**Tabell 5-3:** Sitteplassrekkefølgen for de ulike deltakerne i forsøket. Deltakerne startet ved sitteplass 1 og avsluttet ved sitteplass 3

Tabell 5-4 viser hvordan rekkefølgen på lydeksempelene/oppgavene ble forskjøvet for å unngå at alle deltakere fikk presentert musikken i samme rekkefølge. Rekkefølge A ble brukt på deltaker 1, 4, 7, rekkefølge B ble brukt på deltaker 2, 5, 8, og rekkefølge C på deltaker 3 og 6. I grensesnittet for rekkefølge B og C var oppgavenummereringen som ble vist til deltakeren tilpasset avspillingsrekkefølgen. For å unngå å forvirre leseren, gjenspeiler ikke tabell 5-4 dette.

Rekkefølge A	Lydeksempel	Rekkefølge B	Lydeksempel	Rekkefølge C	Lydeksempel
Oppgave 1	She's got u1	Oppgave 2	Blackbird u1	Oppgave 3	Blackbird u2
Oppgave 2	Blackbird u1	Oppgave 3	Blackbird u2	Oppgave 4	She's got u2
Oppgave 3	Blackbird u2	Oppgave 4	She's got u2	Oppgave 5	Here There
Oppgave 4	She's got u2	Oppgave 5	Here There	Oppgave 1	She's got u1
Oppgave 5	Here There	Oppgave 1	She's got u1	Oppgave 2	Blackbird u1

**Tabell 5-4:** Rekkefølgen på oppgavene/lydeksempelene ble variert for de ulike deltakerne

### 5.5.2 Forsøkspersoner

Som tabell 5-3 viser, deltok åtte personer i forsøket. Utover kjønnsfordelingen, som var jevn, representerte ikke forsøkspersonene et gjennomsnitt av befolkningen. Deltakerne ble rekruttert blant musikalsk- og musikkteknisk utrente personer i undertegnede nærmiljø og arbeidsmiljø, med en aldersmessig sentring rundt 30-tallet.

Til et forsøk av denne art ble det ikke ansett nødvendig å rekruttere bredere i befolkningen, men resultatene av forsøkene må selvsagt ses i lys av det noe snevre utvalget. Alle forsøkspersonene har hørsel på begge ører, men det ble ikke gjennomført noen hørselstest av forsøkspersonene.

### 5.5.3 Instruksjoner til deltakeren

I et forsøk som dette, hvor det fokuseres på ikke-verbal kommunikasjon, er det meget viktig å unngå at det formidles føringer til deltakeren i forkant av gjennomføringen. Grensesnittet til besvarelsen av oppgavene er utformet med dette i tankene. Det er formodentlig intuitivt og enkelt å bruke for deltakerne. Følgende informasjon følte allikevel nødvendig å formidle før start:

*Forløp:* En kort beskrivelse av grensesnittets funksjonalitet, antallet oppgaver og endring av posisjoner underveis ble gitt.

*Fokus:* Deltakeren ble bedt om å kun vurdere opplevd avstand til lydilden(e). I lydeksempel som inneholdt bassgitar, skulle de ikke vurdere denne.

*Vurdering av skalaen:* Deltakernes bruk og vurdering av skalaen var helt opp til dem selv. Det ble det kun formidlet at ”nært” nederst på skalaen betydde ”nært dem selv”.

*Besvarelse ved innsending:* Deltakerne skulle ved sammenligning av de fire versjonene ende opp med et bilde de mente representerte avstandene slik de hadde opplevd det.

*Tilnæringsmåte:* Tilnæringsmåten for sammenligning og vurdering av de fire versjonene var helt opp til deltakeren selv.

*Tidsbruk:* Det ble understreket at deltakeren kunne bruke akkurat så mye eller lite tid de selv følte for.

*Stole på seg selv:* Fordi undertegnede kjenner alle deltakerne ble det til slutt understreket at man skulle besvare oppgaven slik man selv følte det, ikke slik man muligens trodde undertegnede forventet det.

#### **5.5.4 Intervju**

I forlengelsen av det praktiske forsøket, ble det gjennomført et kort intervju med den enkelte deltaker. Hensikten med intervjuet var å få et innblikk i deltakerens opplevelse av forsøket, samt å avdekke hvordan de hadde tenkt i tilnærmingen til oppgavene.

## 6 Analyse

På de kommende sidene vil jeg analysere deltakernes besvarelser i forsøket. Å skulle oversette ikke-verbale uttrykk til meningsfulte ord, må nødvendigvis ha et ambisiøst preg over seg, være seg i kunsten eller vitenskapens navn. Når det ikke-verbale uttrykket i tillegg er formidlet av individer som ikke er trent i genren, øker muligheten for at ukjente faktorer har påvirket uttrykket, som igjen øker muligheten for at oversetteren skyter på feil skive. Det er samtidig umulig å treffe noen skiver som helst hvis man ikke våger å skyte.

Arbeidet er utført med den visshet at analysen av et forsøk som har et såpass lavt presisjonsnivå og flere mulige ukjente feilfaktorer, aldri kan kvalifisere til å være mer enn gode gjetninger. Analysen må også leses med dette i bakhodet.

Opplevelsen av avstand er et av de spatiale trekkene i reprodusert lyd som er vanskeligst å kvantifisere. Utover den basale muligheten til å påvise basale avstandsrelasjoner i dybde dimensjonen er det ingen aspekter som lar seg kvantifisere presist. Dette er årsaken til at graderingen som ble benyttet i forsøket kun formidler sikker informasjon om opplevelsen av lydeksempelene i forhold til hverandre, og at den øvrige bruken av skalaen er overlatt til deltakerens subjektive vurdering. Mulighetene til å tolke deltakernes besvarelser begrenser seg derfor strengt tatt til å peke på tendenser, framfor kvantifiserbare størrelser. På tross av de implisitte begrensningene har undertegnede allikevel dristet seg til å peke på noen sannsynlige forklaringer enkelte steder.

Faren for at det skal oppstå forskningsbias er definitivt til stede i analyser av denne art. Forskningsbias kan manifestere seg i en tilbøyelighet til å fokusere på trekk som underbygger ens egne hypoteser, mens man samtidig unnlater å trekke fram aspekter som peker i negativ.

Undertegnede intensjon er at dette ikke skal skje. Deler av analysen kan derfor bære preg av diskvalifisering av eget arbeid.

I forbindelse med forsøket ble det avholdt et kort intervju med deltakerne i etterkant av besvarelsen av oppgavene. Enkelte steder i analysen vil det bli referert til uttalelser gjort i disse intervjuene.

Analysen vil ikke se detaljert på hver enkelt oppgave, eller hver enkelt besvarelse for den saks skyld. Det synes å være lite hensiktsmessig. For å kunne peke på trender, må man forsøke å holde blikket hevet. Noen detaljer vil allikevel bli trukket fram på steder der det synes relevant.

Romklngen i lydeksemlene var basert på to ulike romklngalgoritmer i TrueVerb (se kapittel 5.2.2). Undertegnede hadde i utgangspunktet tenkt at analysen kunne se på aspekter tilknyttet forskjeller i de to romklngalgorithmsene, men det ble i etterkant vurdert å være vanskelig å gjennomføre, samt litt til side for fokus. I etterkant ser en at det hadde vært mer konsistent å benytte en algoritme på alle lydeksemlene, men det anses ikke å ha betydning for resultatet av forsøket.

Før jeg starter på analysen vil jeg avklare et par sentrale begreper. Selve analysen starter med at jeg setter resultatene opp mot hver av de fem hypotesene som ble formulert i forkant av forsøket. I forlengelsen av hypotesene vil jeg se på andre aspekter som er relevante for forsøket. Jeg begynner med feilprosenten. Den kan være viktig i et forsøk av denne art. Deretter vil jeg i rekkefølge se litt på tidsbruk, bruken av skalaene, vurderingene av gitaren, andre observasjoner og potensielle feilkilder. Jeg avslutter kapittelet med noen oppsummerende betraktninger.

## 6.1 Begreper

Betydningen av ”riktig” og ”galt” (inkludiv synonymene riktig = korrekt og galt = feil) bør være relativt opplagt i denne sammenheng, men for ordens skyld må det sies at klassifisering av en gradering som riktig eller gal defineres av om graderingen gjenspeiler mengden tilførte refleksjoner, stilt opp mot graderingen av de andre versjonene i den enkelte oppgave. Lavere nivå på tidlige refleksjoner må gjenspeiles i lavere verdier på graderingsskalaen for å klassifisere som riktig.

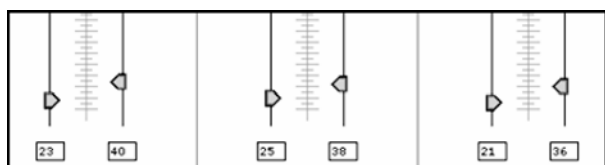
Feilprosenten er regnet ut fra et system hvor hver enkelt gradering gis en poengsum ut fra graderingens posisjon i forhold til de andre graderingene i oppgaven. Utrekningen av poengsummen til den enkelte gradering er etter systemet som vises i tabell 6-1.

Utregning av feilprosent, poengsum til den enkelte gradering	
+1	Graderingen er riktig plassert i forhold til graderingen den sammenlignes med
0	Graderingen er likt gradert som graderingen den sammenlignes med (slingringsmonn +-2)
-1	Graderingen er feil plassert i forhold til graderingen den sammenlignes med

**Tabell 6-1:** Kriterier (poengsummer) for utregning av feilprosent

I oppgaver med 4 versjoner vil den enkelte gradering få en poengsum fra -3 til +3, mens skalaen går fra -2 til +2 i oppgaver med 3 versjoner. Poengsummene er regnet om til prosenter for bedre å synliggjøre resultatene. Null prosent innebærer at alle graderinger står i korrekt posisjon, mens hundre prosent betyr at alle graderinger er feil plassert.

I tabell 6-1 ser man at graderinger innenfor en avstand på pluss/minus 2 anses å være likt gradert. Slingringsmonnet er lagt inn på bakgrunn av intervjuene. Figur 6-1 illustrerer den visuelle likheten i avstander på 2 piksler i grensesnittet.



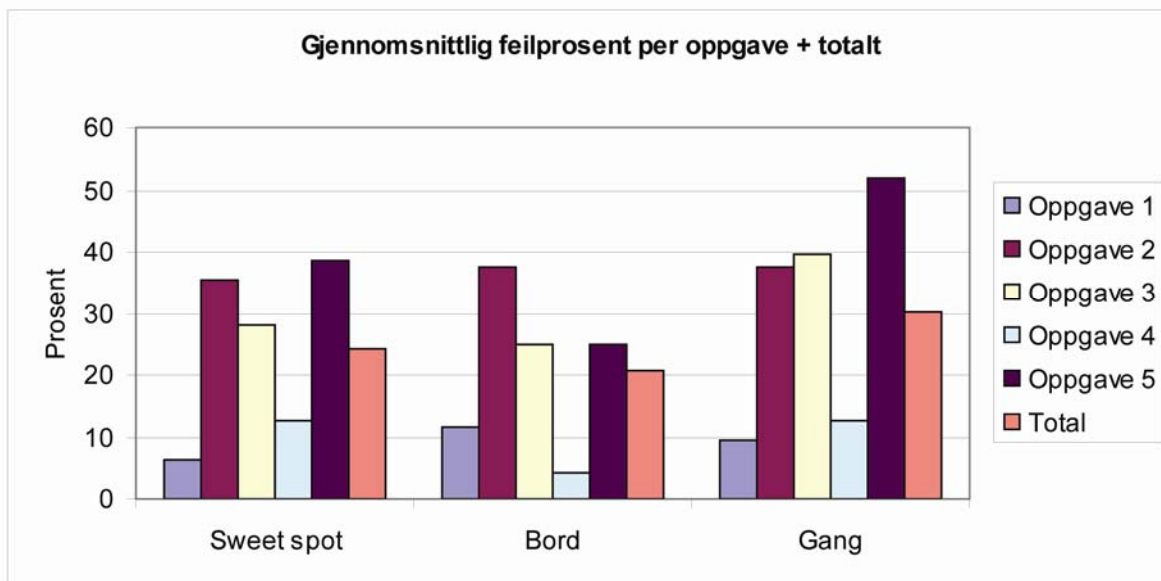
**Figur 6-1:** Graderinger med en avstand på to piksler anses som like (i bildet blir annenhver gradering ansett som lik)

## 6.2 Hypotesene

Før jeg beveger meg inn på hypotesene, bør det nevnes at det allerede ved første blick på grafene fra oppgavebesvarelsene er en oppgave som skiller seg ut fra resten. Oppgavene er beskrevet i tabell 5-2. Hadde alle besvarelsene hatt samme preg som oppgave 1, som tilsynelatende bekrefter flere av forfatterens hypoteser, ville resultatet vært akkurat slik undertegnede antok i forkant av forsøket. Det ble ikke akkurat slik.

### 6.2.1 Hypotese 1

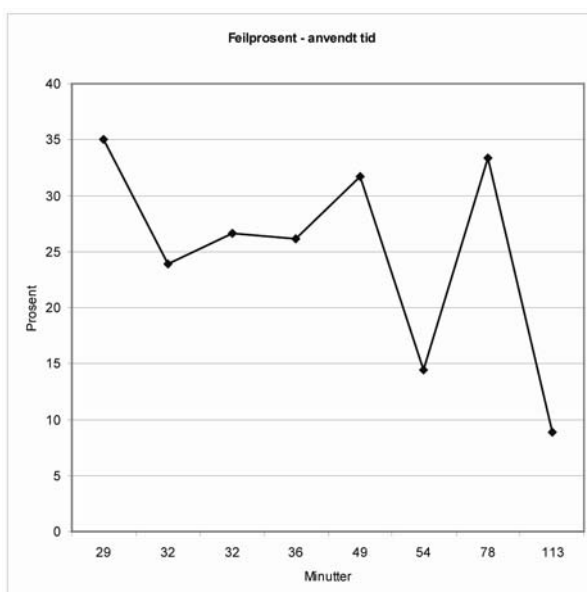
*”Feilprosenten vil være høyere ved posisjon 2 og 3 enn posisjon 1, og antageligvis høyere ved posisjon 3 enn posisjon 2.”* Sammenligner vi den gjennomsnittlige feilprosenten for alle besvarelser i hver av posisjonene ser vi at første del av hypotesen ikke blir bekreftet, mens andre del av hypotesen blir det. Feilprosenten var 24,1 %, 20,5 % og 30,2 % i henholdsvis posisjon 1, 2 og 3. Vi ser at en oppbryting per oppgave (se figur 6-2) heller ikke bekrefter hypotesen.



**Figur 6-2:** Søylen viser gjennomsnittlig feilprosent i hver oppgave fordelt på posisjoner. Søylen helt til høyre ved hver posisjon viser gjennomsnittet for alle oppgaver samlet

## 6.2.2 Hypotese 2

”Feilprosenten vil gå ned ved mer anvendt tid.” Hypotese 2 blir ikke bekreftet (se figur 6-3), men den overordnede trenden kan tyde på at hypotesen har en viss berettigelse. Det er påtagelig at den personen med desidert høyest anvendt tid, også er den personen med desidert lavest feilprosent.



**Figur 6-3:** Gjennomsnittlig feilprosent for den enkelte deltaker fordelt på anvendt tid



### 6.2.3 Hypotese 3

”Det vil være høyere feilprosent mot lavere nivåer av tidlige refleksjoner.” På grunn av utelatelsen av en versjon i oppgave 2 og 4, har det bare vært mulig å teste hypotesen opp mot oppgave 1, 3 og 5. Tabell 6-2 viser feilprosent for den enkelte versjon, både det totale gjennomsnittet og fordelt på posisjoner. Fra versjon 1 til 3 stiger feilprosenten, men versjon 4 har lavere feilprosent enn versjon 2 og 3 i alle posisjoner. Et resultat som ikke gjenspeiles i tabellen nedenfor, men som ville trukket versjon 4 enda lavere, er at versjon 4 i oppgave 4 har null i feilprosent. Hypotese 3 er dermed ikke bekreftet.

	Versjon 1	Versjon 2	Versjon 3	Versjon 4
Samlet gjennomsnitt, alle posisjoner	17,6 %	29,4 %	31,3 %	26,4 %
Posisjon 1 (sweet spot)	14,6 %	28,5 %	31,3 %	22,9 %
Posisjon 2 (bord)	12,5 %	23,6 %	25 %	20,8 %
Posisjon 3 (gang)	25,7 %	36,1 %	37,5 %	35,4 %

**Tabell 6-2:** Gjennomsnitt feilprosent fordelt på versjoner (De to like verdiene er ikke en tastefeil)

### 6.2.4 Hypotese 4

”Man vil se en utflating i graderingen mellom versjon 3 og 4.” Det er vanskelig å bekrefte denne hypotesen. Generelt spriker besvarelsene for mye til å kunne trekke slutninger av denne art. Sammenligner vi kun helt korrekte besvarelser i oppgaver med 4 versjoner (de er relativt få utover oppgave 1), bekrefte hypotesen, men ikke med overbevisende tall. Det er definitivt noe tendensiøst å filtrere bort ukorrekte besvarelser, men det er en interessant øvelse. Tabell 6-3 viser gjennomsnittlig avstand mellom graderingene av de fire versjonene. Tabellen er brutt ned på den enkelte posisjon, med et totalgjennomsnitt nederst. Vi ser at posisjon 1 utmerker seg til fordel for hypotesen, mens tallene for posisjon 2 og 3 diskvalifiserer den. Det er interessant å merke seg denne forskjellen mellom posisjonene, da det muligens kan anses å belyse vår evne til å oppleve nyanser i og utenfor sweet spot.

	Versjon 1 til 2	Versjon 2 til 3	Versjon 3 til 4
Posisjon 1 (sweet)	24	24	13
Posisjon 2 (bord)	18	12	15
Posisjon 3 (gang)	22	20	23
Gjennomsnitt total	21	19	16

**Tabell 6-3:** Gjennomsnittlig avstand i graderinger fra en versjon til neste i oppgaver med 4 versjoner

### 6.2.5 Hypotese 5

*”Mer komplekst lydbilde og færre transienter vil gjenspeiles i større sprik i graderinger samt høyere feilprosent.”* En visuell sammenligning av graderingene i oppgavene viser at denne hypotesen må sies å bekreftes. Det visuelle inntrykket bekreftes ytterligere av feilprosenten (se figur 6-2). Oppgave 1 og 4 har langt flere korrekte besvarelser enn de andre oppgavene. Dette er de to oppgavene som kun består av a cappella kvinnestemme. Begge lydeksemlene er rike på transienter og luft. Oppgave 2 består også kun av vokal, men strofen har en helt annen karakter enn kvinnestemmen. Mot slutten av strofen får stemmen en litt mer dynamisk, konsonantrikt karakter, men da forsøksgrensesnittet kun innbød til grundig sammenligning i starten av frasene var det i hovedsak der den reelle sammenligningen foregikk.

Gitaren som akkompagnerer vokalen i oppgave 3 og 5, samt bassen i oppgave 5, bidrar til et langt mer komplekst lydbilde i disse to oppgavene. Lydbildets kompleksitet gjør avstandsbedømmingen vanskeligere for lytteren, noe som reflekteres i den høyere feilprosenten.

### 6.3 Feilprosent

Man kan ikke unngå å legge merke til at feilprosenten er relativt høy i noen av oppgavene, mens den samtidig er gledelig lav i andre. I de følgende avsnitt skal jeg se på ulike forhold som kan belyse feilprosenten.

Det er, som beskrevet under hypotese 5, et godt samsvar mellom feilprosent og lydbildets kompleksitet/transientrikdom. En del av feilprosenten kan derfor tilskrives faktorer som gjør avstandsbedømmingen vanskeligere for lytteren. Et argument som taler mot denne koplingen er at man muligens skulle anta at vanskeligheter i avstandsbedømmingen ville manifestere seg i en lik gradering av eksemplene. Jeg kommer tilbake til lik gradering nedenfor, men med unntak av en deltaker, samt en oppgave/posisjon, er bruken av lik gradering forholdsvis lav. Man kan derfor spekulere i om en mindre tydelig avstandsindikasjon vil gjøre deltakeren mer tilbøyelig til å ”gjette”.

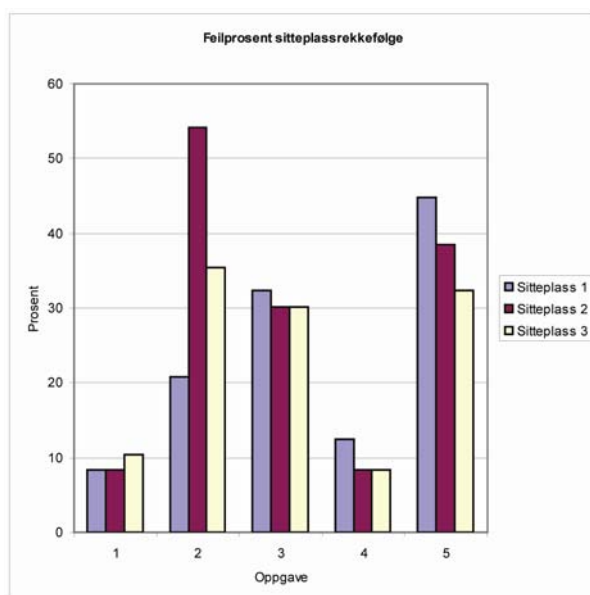
Flere av deltakerne fortalte under intervjuet at de i oppgavene med gitar fant det vanskelig å formidle opplevelsen av forholdet mellom gitar og vokal i den enkelte versjon, og på samme tid formidle forholdet mellom de ulike versjonene i en oppgave. Man må derfor anta at dette har medvirket til høyere feilprosent.

Tabell 6-2 viser at versjon 1 har en lavere feilprosent enn de andre versjonene. En kan anta at årsaken ligger i det noe overdrevne nivået på de tidlige refleksjonene i versjon 1. Lydnivået har altså gjort det enklere for lytteren å identifisere avstanden.

Feilprosenten må også ses i lys av at deltakerne var uskolerte lyttere. I neste avsnitt vil vi se at deltakerne brukte lengst tid ved sitteclass 1, altså posisjonen hvor de besvarte de fem første oppgavene, og at dette kan indikere at lytterne har lært underveis. I lys av det, er det naturlig å spørre om feilprosenten gjenspeiler seg i sitteclassrekkefølgen. Tabell 6-4 viser den samlede feilprosenten fordelt på sitteclassrekkefølge, mens figur 6-4 viser det samme fordelt på den enkelte oppgave. Vi ser først og fremst at det ikke er noe konsekvent forhold mellom feilprosent og sitteclassrekkefølge. Oppgave 2 trekker åpenbart kraftig i negativ retning. Samtidig er det interessant å merke seg at feilprosenten for oppgave 5, som er den oppgaven med det mest komplekse lydbildet med hensyn til antall instrumenter (vokal, gitar og bass), avtar etter hvert som forsøket skrider frem.

Sitteclass 1	23,8 %
Sitteclass 2	27,9 %
Sitteclass 3	23,3 %

**Tabell 6-4:** Gjennomsnittlig feilprosent fordelt på sitteclassrekkefølge



**Figur 6-4:** Feilprosent fordelt på sitteclassrekkefølge. Sitteclass 1 er posisjonen der deltakeren startet forsøket – Sitteclass 3 er posisjonen der deltakeren avsluttet forsøket. Rekkefølgen på posisjoner var forskjellig mellom deltakerne

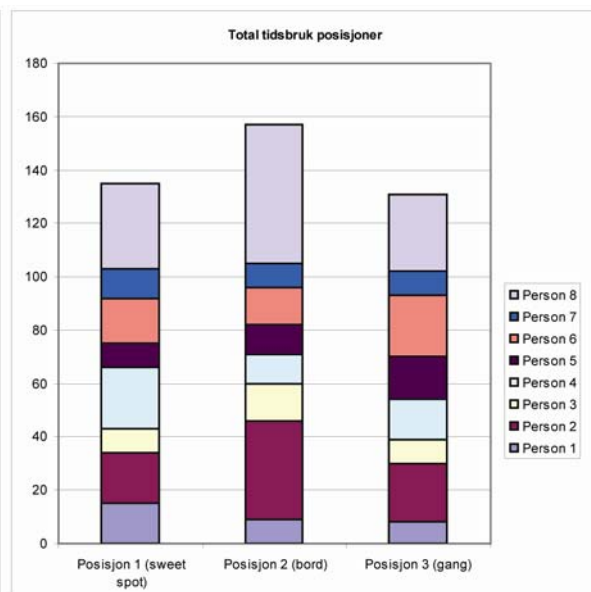
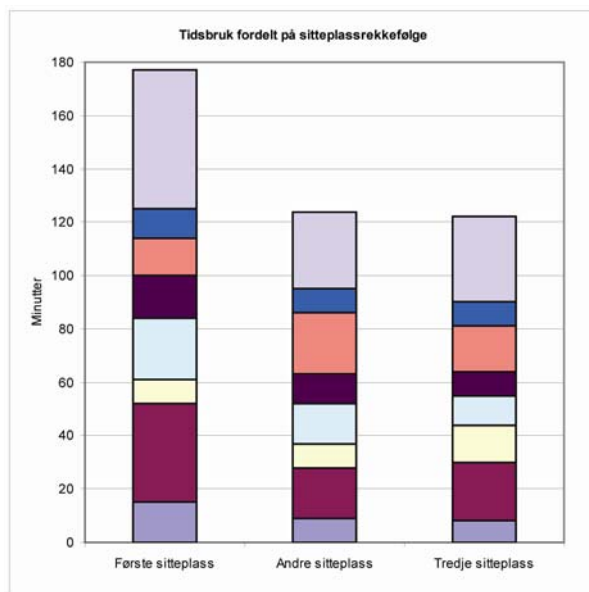
Feilprosenten må også ses i lys av muligheten for at deltakerne har vurdert andre aspekter i lyden enn den avstandsindikasjonen som følger med de tidlige refleksjonene. Under intervjuet med en av deltakerne med høyest feilprosent, fortalte denne om sitt forsøk på å identifisere avstandsopplevelsen ut fra en assosiasjon til ulike rom.

Til slutt skal man ikke glemme at selv om feilprosent og sprik i besvarelsene har en negativ klang, illustrerer de også utfordringene konsumenter vil møte hvis de i sin lytting forsøker å gripe spatiale trekk i reprodusert lyd.

## 6.4 Tidsbruk

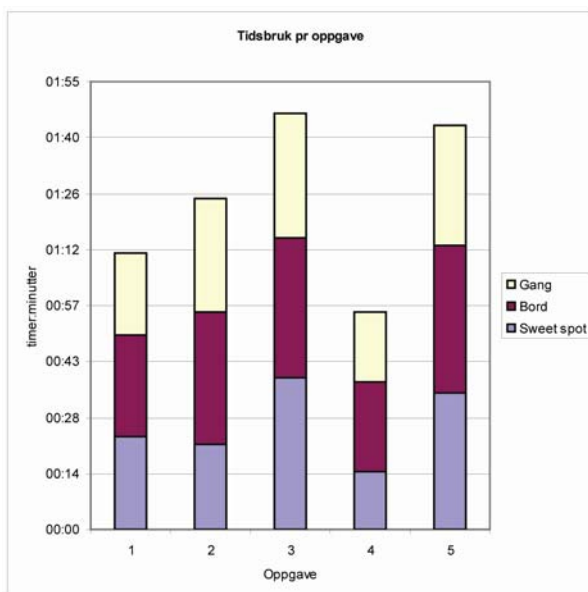
Det er flere sider ved deltakernes tidsbruk som er interessant å se på. Ser man på deltakernes tidsbruk i forhold til sitteplassrekkefølge (se figur 6-5), minker den totale tidsbruken per posisjon etter hvert som forsøket skrider fram. Dette tyder på at deltakerne har lært underveis. Eventuell læring må ses i sammenheng med at dette var en ny situasjon for deltakerne. Man ser samtidig at den gradvis nedgående tidsbruken ikke gjelder for alle deltakere.

Sammenligner vi tidsbruken mellom de ulike posisjonene (se figur 6-6) er det ingen klar sammenheng. Posisjonen ved bordet er noe høyere, men det må hovedsakelig tilskrives tidsbruken til en enkelt deltaker.

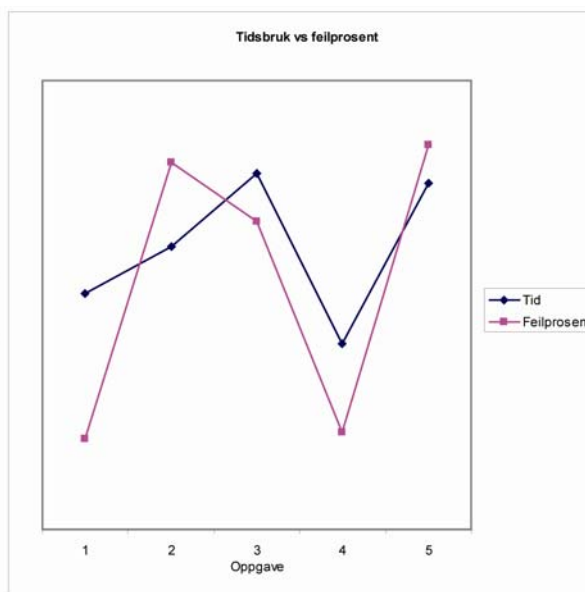


**Figur 6-5:** Anvendt tid for delt på sitteplassrekkefølge **Figur 6-6:** Anvendt tid fordelt på posisjoner

Figur 6-7 viser den totale tidsbruken per oppgave. Vi ser her en tydelig kopling mellom oppgavens kompleksitet og den anvendte tiden. Figur 6-8 viser det klare sammenfallet mellom tidsbruk og feilprosent. Den klare koplingen er en klar indikasjon på at lydbildets kompleksitet er langt mer avgjørende for feilprosenten enn anvendt tid i seg selv.



**Figur 6-7:** Tidsbruk fordelt på oppgaver



**Figur 6-8:** Tidsbruk versus feilprosent fordelt på oppgaver

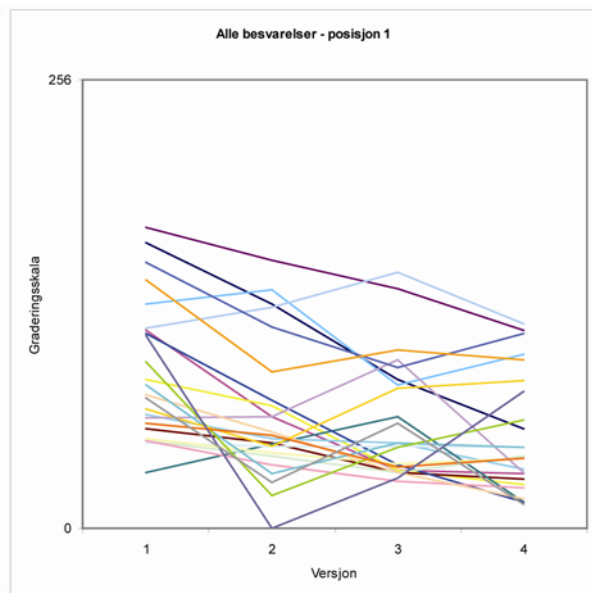
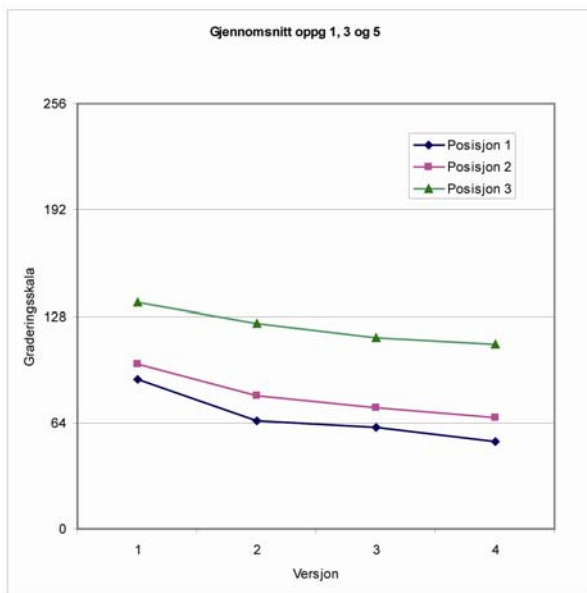
Som en liten kuriositet på slutten av dette avsnittet, viser det seg at kvinnene i gjennomsnitt brukte nesten dobbelt så lang tid som mennene på å løse oppgavene, mens feilprosenten mellom kjønnene hadde en avstand på to promille.

## 6.5 Bruk av skalaen

I dette avsnittet skal jeg se på noen aspekter knyttet til deltakernes bruk av graderingsskalaen som ble benyttet i grensenettet til lytteforsøket. Utover den grunnleggende instruksjonen om at opplevd avstand skulle gjenspeiles i graderingen, var det opp til den enkelte deltaker å avgjøre hvordan han/hun benyttet graderingsskalaen.

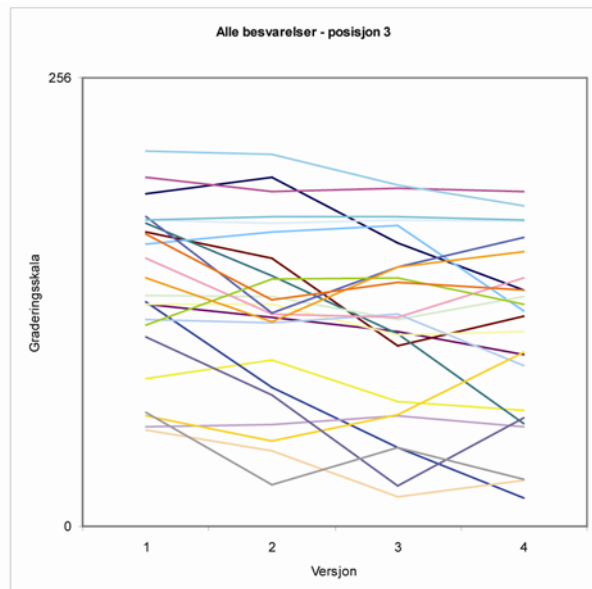
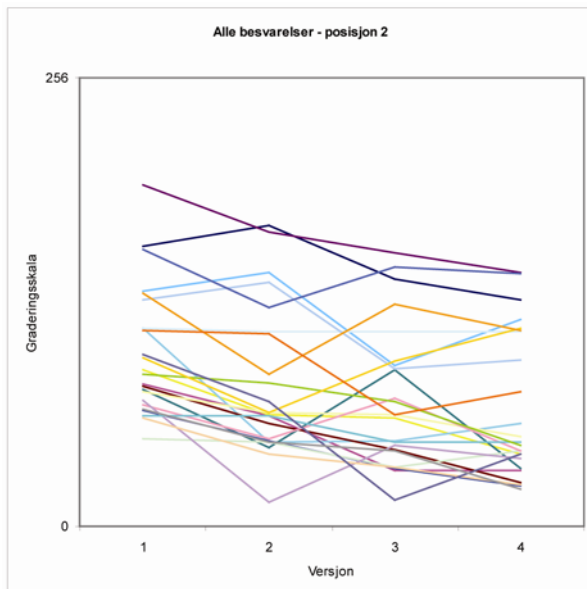
Figur 6-9 til 6-12 viser en samling av alle besvarelser i oppgavene med fire versjoner. En ser på gjennomsnittet at deltakerne generelt tenderer til å gradere i nedre del av skalaen. I grafene som viser alle graderingene, ser en hvordan spredningen øker ved økt avstand til høyttalerne. Forholdet mellom gjennomsnittsverdiene gjenspeiler også posisjonenes avstand til høyttalerne, men ikke proporsjonalt med den reelle avstanden. Avstanden mellom lytteren og nærmeste høyttaler i posisjon 1 til 3, var henholdsvis 125 cm, 400 cm og 650 cm. Avstanden i gjennomsnittet for posisjon 2 og 3 er således langt større enn de fysiske centimeterne skulle tilsi, sammenlignet med avstanden mellom posisjon 1 og 2. Man kan spekulere i en rekke potensielle faktorer som kan spille en rolle for graderingen, men det er allikevel to trekk ved som skiller seg ut. Først og fremst var posisjon 3 plassert i et annet rom enn høyttalerne, og i tillegg hadde man ikke fri sikt til høyttalerne.

En legger merke til at gjennomsnittsgrafene er korrekte med hensyn til innbyrdes posisjon for versjonene, samt at gjennomsnittet taler til fordel for hypotese 4. Vi ser samtidig at det er en større avstand mellom minimums- og maksimumsverdien for den enkelte posisjon. Avstanden er 37, 32 og 25 for henholdsvis posisjon 1, 2 og 3. Det indikerer en utflating i avstandsopplevelsen som følge av den økte avstanden.



**Figur 6-9:** Gjennomsnittsverdien av alle deltakernes besvarelser i oppgave 1, 3 og 5

**Figur 6-10:** Alle deltakernes besvarelser i posisjon 1, oppgave 1, 3 og 5



**Figur 6-11:** Alle deltakernes besvarelser i posisjon 2, oppgave 1, 3 og 5

**Figur 6-12:** Alle deltakernes besvarelser i posisjon 3, oppgave 1, 3 og 5

Et gjennomgående trekk i besvarelsene er deltakernes tendens til å plassere den enkelte gradering like ved hjelpelinjene lagt inn for hver 16. piksel i den 256 piksel høye graderingsskalaen. Hjelpelinjene kan i så måte sies å ha medvirket til å senke den reelle oppløsningen på graderingsskalaen. Det er allikevel tvilsomt at det har hatt noen avgjørende betydning for forsøket.

Tendensen til å plassere graderingene i nedre del av skalaen kan ha blitt påvirket av det faktum at den økte mengden tidlige refleksjoner i posisjon 2 og 3, som begge i tillegg befinner seg off-axis i forhold til lyd-kilden, kan ha medført en økning i opplevd lydstyrke. De tidlige refleksjonene vil mest sannsynlig gi likt utslag mellom de ulike versjonene, men mellom posisjoner kan man ha opplevd en økning i lydstyrke. Lydtrykket fra høyttalerne ble justert mellom posisjoner, men trenger ikke å ha gjenspeilt den opplevde lydstyrken.

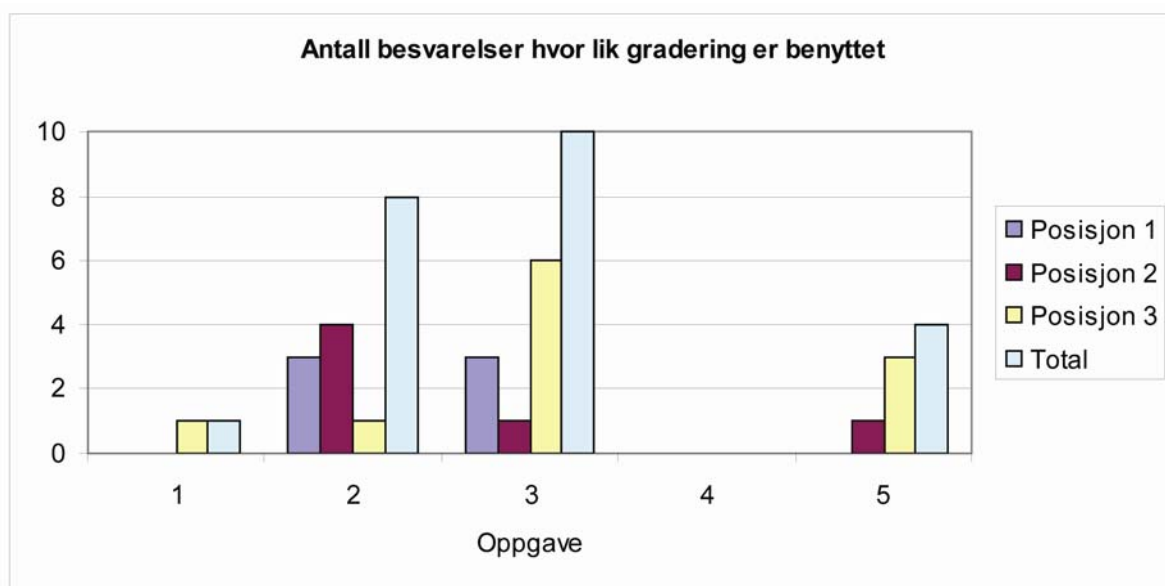
Ordet ”fjernt” var plassert like over graderingsskalaen for å unngå at graderingsskalaen ble benyttet opp ned. I samme hensikt var ordet ”nært” plassert i underkant av skalaen. Ordet fjernt kan gi assosiasjoner til noe som er plassert langt borte. Avstandsendingene som følge av de tidlige refleksjonene alene var ikke i stand til å simulere noe langt borte. Lydbildet gav således et generelt ”nært” inntrykk, uavhengig av versjonene.

### **6.5.1 Lik gradering**

Det totale antallet oppgavebesvarelser i forsøket var 120, fordelt på 40 besvarelser i hver posisjon. Av de 120 besvarelsene var det 23 som inneholdt en eller flere like graderinger. Av de 23 besvarelsene kom 9 av dem fra en enkelt deltaker. Tabell 6-5 viser fordelingen av de ulike kategoriene oppgavebesvarelser med lik gradering faller inn under. Figur 6-13 viser fordelingen mellom posisjoner og oppgaver. Vi ser at det er en overvekt i bruken av lik gradering ved posisjon 3. Det er også et sterkt sammenfall mellom lydbildets kompleksitet, som beskrevet i hypotese 5, og bruken av lik gradering. Et interessant trekk å merke seg skjer i oppgave 3, posisjon 3 (gang). Der har 6 av 8 deltakere gradert en eller flere versjoner likt.

Kategorier oppgavebesvarelser med lik gradering	Posisjon 1	Posisjon 2	Posisjon 3
Oppgaver der to tilstøtende versjoner er gradert likt	4	1	4
Oppgaver der to ikke tilstøtende versjoner er gradert likt	1	1	1
Oppgaver der to og to tilstøtende versjoner er gradert likt		1	1
Oppgaver der tre tilstøtende versjoner er gradert likt			1
Oppgaver der tre versjoner (to tilstøtende og en ikke tilstøtende) er gradert likt			2
Oppgaver der alle tre versjonene er gradert likt		2	
Oppgaver der alle fire versjonene er gradert likt	1	1	2
Totalt	6	6	11

**Tabell 6-5:** Ulike varianter av lik gradering



**Figur 6-13:** Antall besvarelser hvor deltakeren har gradert to eller flere versjoner likt. I oppgave 1 var det ingen like graderinger ved posisjon 1 og 2. I oppgave 4 var det ingen like graderinger

## 6.6 Gitaren

I musikk eksempelen til oppgave 3 og 5 er vokalen akkompagnert av gitar. Oppgave 5 inneholder i tillegg bassgitar. I disse to oppgavene skulle deltakerne gradere både vokal og gitar, bassen skulle ikke vurderes. Gitaren i de to lydseksemplene var ikke tilført noen form for romklang. Den ble dermed reproduert likt mellom versjonene. Gitaren er følgelig ikke noe primært mål for analysen, men kan med sitt statistiske preg synliggjøre interessante aspekter ved deltakernes vurderinger.

Studerer man grafene for gitaren (se Appendiks A) ser man tydelig at graderingen i en rekke besvarelser er langt fra flat. Generelt synes oppgave 3 å ha et noe flatere preg enn oppgave 5, og posisjon 3 i begge oppgavene synes å ha et flatere preg enn de to andre posisjonene. En enkel addering av differansen mellom maksimums- og



minimumsgraderingene for de enkelte besvarelsene bekrefter påstandene i forrige setning. I oppgave 3 er verdiene 162, 194, 150 for henholdsvis posisjon 1, 2 og 3. I oppgave 5 er verdiene 285, 236 og 188. Av de 48 besvarelsene hvor gitarens plassering er vurdert, inneholder 31 besvarelser lik gradering av to eller flere versjoner.

Den ujevne graderingen av gitaren belyser godt den uskolerte lytterens vanskeligheter ved å vurdere spatiale trekk i reprodusert musikk. Flere av forsøkspersonene nevnte uoppfordret i intervjuet at de hadde problemer med å gradere gitaren. Flere sa at de opplevde den som lik, men at de allikevel hadde gradert den forskjellig. Som nevnt tidligere, sa også flere av deltakerne at de hadde problemer med å formidle et bilde hvor både forholdet mellom vokal og gitar i hver enkelt versjon og forholdet mellom de ulike versjonene ble ivaretatt. En kan argumentere at den utrente lytteren sannsynligvis har bedre forutsetninger for å vurdere spatiale trekk knyttet til den primære meningsbæreren i musikken, framfor den akkompagnerende gitaren. Samt at den primære meningsbæreren naturlig vil legge krav på mest oppmerksomhet. Det er dermed sannsynlig at gitaren er den mest skadelidende i dilemmaet flere av deltakerne følte seg stilt overfor.

Det noe flatere preget i oppgave 3 sammenlignet med oppgave 5, kan koples til den økte kompleksiteten som følge av et tredje instrument i musikken til oppgave 5. Det er vanskeligere å finne en åpenbar grunn til at posisjon 3 utmerker seg i positiv retning. Man må anta at forklaringen er å finne i en kombinasjon av den økte avstanden, refleksjonene i det separate (tilkoplete) rommet og den obstruerte banen til den direkte lyden, men det er problematisk å peke på noe mer konkret enn det.

## 6.7 Noen observasjoner

Før jeg beveger meg over til mulige feilkilder i forsøket, skal jeg presentere to observasjoner som er påtagelige når man studerer grafene visuelt.

I oppgave 5 skiller versjon 1 seg ut fra de andre versjonene. Dette kan tolkes i retning av at det kun var et overdrevent nivå på de tidlige refleksjonene (+3dB) som tydelig formidlet endring i avstand. Hvis dette stemmer er det i tråd med hypotese 4 og 5, og belyser de utfordringer lytteren stilles overfor ved et mer komplekst lydbilde.

I oppgave 4 er feilprosenten markant lavere ved posisjon 2 (bord) enn ved posisjon 1 (sweet spot). Ser man nærmere på feilbesvarelsene ved posisjon 1, har alle plassert versjonen med -3 dB, som opprinnelig var versjon 3, lenger borte enn versjon 1 (+6dB). Når feilbesvarelsene er så unisone, øker muligheten for at det er en faktor som har påvirket vurderingen. En mulig forklaring i denne oppgaven er at når de tilførte refleksjonene er 6 dB over den direkte lyden,

gir dette et "unaturlig" styrkeforhold der refleksjonene er veldig tydelige i seg selv. For de deltakerne det gjelder, er det mulig det kraftige nivået på de tidlige refleksjonene ikke er opplevd som en økning i avstanden, men mer som en endring av lydkildens klanglige karakteristikk. Lydstyrken på de tidlige refleksjonene i +6 dB-versjonene er så overdrevne, at innspillinger med denne ratioen neppe har sett en hitliste i sitt liv.

## 6.8 Mulige feilkilder

At dette forsøket beveger seg i noe upløyd mark med hensyn til både lytteposisjon og deltakere, bidrar til å øke antallet potensielle feilkilder. De ukontrollerte omgivelsene hvor forsøket fant sted, samt undertegneds manglende erfaring på området, bidrar heller ikke i positiv retning. Enkelte mulige feilkilder er allerede beskrevet tidligere i analysen. Nedenfor følger en beskrivelse av flere potensielle feilkilder. Opplistingen er antakeligvis ikke utfyllende.

Det er en åpenbar mulighet for at de uskolerte deltakerne har fokusert på andre klanglige attributter enn den avstandsindikasjonen som følger av de tidlige refleksjonene. Det har allerede vært nevnt to steder lenger oppe, at en deltaker hadde forsøkt å visualisere hvilket rom lydkilden i befant seg i. Deltakeren meddelte videre en opplevelse ulike rom innenfor den enkelte oppgave. Assosiasjonen til ulike rom er en helt legitim opplevelse ut fra de ulike nivåene på refleksjonene, men det åpner samtidig for at graderingen har sitt utspring i andre klanglige attributter. Det ble ikke gjort noe forsøk på å identifisere hvilke oppgaver og versjoner det var snakk om. Det kan spekuleres om ikke +6 dB- og +3 dB-versjonene er involvert. Da disse med sin overdrevne tydelighet også innebærer en endring av lydkildens klanglige karakteristikk. En annen deltaker opplevde oppgavene som vanskelige, og fortalte i etterkant om å ha måttet "lete" etter hva som skulle vurderes.

I forlengelsen av forrige avsnitt, må det nevnes at alle deltakerne er bekjente av undertegnede. Alle deltakerne visste således at romklang er fokus for oppgaven. Det bør i utgangspunktet være legitimt å benytte bekjente i lytteforsøk av denne art, men det kan tenkes at kombinasjonen av at deltakerne er uskolerte og kjennskapet til undertegneds fokus har gjort dem mer tilbøyelige til å vurdere andre aspekter enn de skulle.

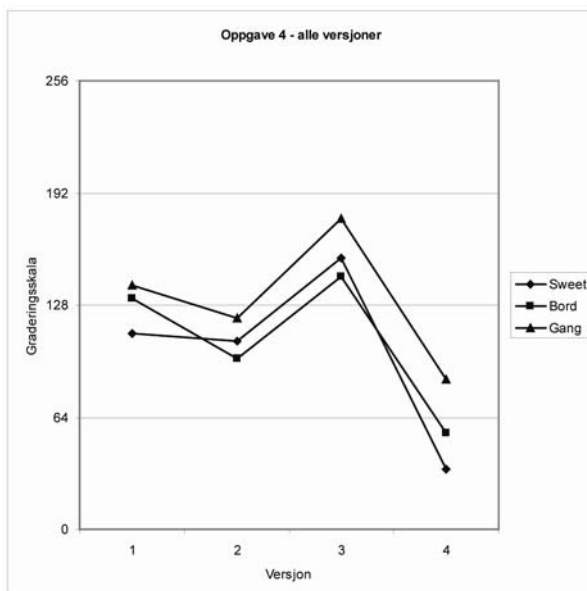
Selv om det ble arbeidet grundig med nivåjustering, er det allikevel mulighet for at nivåene ikke var helt like. Dette kunne antakeligvis vært kontrollert enda bedre ved å benytte flere lyttere under nivåjusteringen.

Det var åtte personer som deltok i forsøket. Det lave antallet er i seg selv en potensiell feilkilde. Man kan anta at enkelte trekk og usikkerhetsfaktorer kunne vært ryddet av veien ved

et høyere antall deltakere. Et høyere antall deltakere ville trukket forsøket mer i retning av et kvantitativt forsøk, som analysen åpenbart tenderer til å være.

Det må nevnes at det generelle spørsmålet om hvorfor noen svar er feil mens tendensen peker i riktig retning ikke utelukkende kan tilskrives deltakernes uskolerthet. Man kan ikke avskrive at andre potensielle feilkilder har spilt en rolle. Avslutningsvis er det viktig å nevne at også de positive resultatene må ses i lys av mulige feilkilder. Det er usannsynlig at det bare er de ukorrekte graderingene som har vært påvirket av andre faktorer, eller eventuelt har vært gjenstand for gjetning. Korrekte graderinger kan med andre ord ha havnet i riktig posisjon på bakgrunn av helt ”feilaktige” vurderinger.

Som en liten kuriositet helt til slutt viser figur 6-14 konsekvensen av nivåfeilen som resulterte i at to av versjonene i forsøket, en i oppgave 2 og en i oppgave 4, måtte utelukkes fra resultatene.



**Figur 6-14:** Gjennomsnittsverdier av deltakernes besvarelser i oppgave 4. Versjon 3 ble sett bort fra i analysen på grunn av nivåfeil i avspillingen

## 6.9 Avsluttende betraktninger

Vi har sett at feilprosenten er varierende og at kompleksiteten i materialet har størst betydning for resultatet. Man kan dermed anta at konsumenter relativt godt evner å oppfatte avstandsendringer som følge av tidlige refleksjoner, forutsatt at lydbildet ikke er for komplekst. Øker kompleksiteten i lydbildet vil konsumentens evner utfordres i større grad.

Den høye feilprosenten i oppgave 3 og 5 (se figur 6-2) må til dels ses i lys av at deltakerne her skulle løse to oppgaver. De skulle vurdere avstanden til både vokal og gitar i den enkelte versjon, samtidig som de skulle vurdere de ulike versjonene opp mot hverandre.

Det ble gitt tydelige tilbakemeldinger på at dette kompliserte vurderingen for flere av deltakerne. Den lite flate graderingen av gitaren kan anses å belyse utfordringen deltakerne ble stilt overfor. Det er vanskelig å si i hvilken grad den økte vanskelighetsgraden som følge av å skulle gradere to instrumenter har bidratt til å øke feilprosenten. Sett i sammenheng med oppgave 2 som hadde høy feilprosent, men ikke gitar, kan man anta at den økte vanskelighetsgraden ikke er nok til å diskvalifisere hypotese 5.

Hypotese 1 som sier at feilprosenten skal øke i posisjon 2 og 3 blir ikke innfridd, men siste del av hypotesen blir det. Siste del sier at feilprosenten antakeligvis vil være høyere ved posisjon 3 enn posisjon 2. Resultatet kan tolkes i retning av at hypotesen har en berettigelse, men at overlevelsen til spatiale trekk ved økt avstand til høyttalerne, er langt fra så forutsigbare som hypotesen fremstiller det. Hypotese 3 sier at feilprosenten vil være høyest mot lavere nivåer av tidlige refleksjoner. Denne hypotesen blir heller ikke bekreftet, med unntak av versjonen med høyest nivå (versjon 1) (se tabell 6-2).

Ved å sammenligne alle graderinger i den enkelte posisjon lagt over hverandre (se figur 6-10 til 6-12) viser bruken av graderingsskalaen økt spredning ved økt avstand til høyttalerne. Både gjennomsnittsgrafene og de samlede grafene for hver posisjon, viser en tydelig trend som peker i retning av at deltakerne samlet har evnet å oppfatte ulikhetene i avstand som følge av de kunstig genererte tidlige refleksjonene. Vi ser samtidig at denne trenden flater ut ved posisjon 3. Videre taler det visuelle inntrykket til fordel for hypotese 4 som antar en utflating mot versjon 3 og 4. Det visuelle inntrykket står i og for seg ikke i opposisjon til de tallbaserte resultatene, men øker tilbøyeligheten til å si at resultatene trekker i retning av det som var forventet. Foregående setning kan tolkes som en antydning om at visuelle vurderinger er bedre egnet til synliggjøring av resultater når tendenser og trender er målet. Men man kan like godt hevde det motsatte.

Den økende spredningen i bruken av graderingsskalaen fra posisjon 1 til 3 (6-10 til 6-12) sier oss lite om forskjellene i det innspilte materialet, men synliggjør deltakernes varierende opplevelse av avstanden mellom seg selv og den reproduserte musikken.

Bruken av like graderinger var nesten dobbelt så høy ved posisjon 3 som ved posisjon 1 og 2. Det forsterker inntrykket av at deltakernes opplevelse i posisjon 3 skiller seg ut fra de to andre posisjonene. Vi ser også at bruken av like graderinger følger oppgavens kompleksitet.

Hypotese 2 som antok lavere feilprosent ved høyere tidsbruk, ble som kjent ikke bekreftet. Trenden peker allikevel i riktig retning. Et aspekt som forkludrer etterforskningen av hypotese 2, er at en må anta at deltakerne hadde ulike forutsetninger for å løse oppgavene i

utgangspunktet. Det lave antallet deltakere i forsøket gjør det vanskelig å konkludere. Det hadde generelt vært en fordel med et langt høyere antall deltakere for å foreta gode statistiske analyser.

Tidsbruksstatistikken viser at deltakerne brukte lengst tid ved første sitteplass (se figur 6-5), mens feilprosenten ikke fulgte samme utvikling (se tabell 6-4). Dette kan tolkes til fordel for hypotese 2. Vi så også at tidsbruken fulgte kompleksiteten på oppgavene (se figur 6-7).

Gitaren har vært viet lite oppmerksomhet, men i tillegg til å ha bidratt til å komplisere oppgave 3 og 5 for deltakerne, kan den anses å bidra til å belyse de utfordringer konsumenter møter hvis de forsøker å vurdere spatiale trekk i reprodusert musikk.

Til slutt har vi sett at det knytter seg en rekke potensielle feilkilder til resultatet i forsøket. En hver uttalelse og antagelse må derfor tas med en stor klype salt.

## 7 Konklusjon

Etter noen og nitti sider med hovedoppgavestoff gjenstår nå bare to små spørsmål. Hva har vi lært, og hvordan kan vi lære mer? Det er naturlig å starte med å se på hva vi har lært. Problemstillingen for oppgaven har hatt som mål å belyse hvordan konsumenter oppfatter spatiale trekk i populærmusikkproduksjoner reprodusert i stereo. Problemstillingen har i utgangspunktet et stort nedslagsfelt. Som et tiltak for gjøre problemstillingen håndterbar, ble det valgt å fokusere på romklangens tidlige refleksjoner. Det ble også valgt å fokusere på ikke-optimale lyttesituasjoner. Fokuset på det ikke-optimale fulgte som en naturlig konsekvens av ønsket om å belyse musikkopplevelsen til konsumenten framfor den skolerte lytteren. Førstnevnte gruppe lytter sjelden til musikk under optimale forhold. Fokuset på konsumenten og de ikke-optimale lyttesituasjonene plasserer denne oppgaven noe til side for musikkvitenskapens tradisjonelle innfallsvinkel når den skal se på spatiale trekk i reprodusert musikk. Musikkvitenskapen har, i likhet med en rekke andre vitenskaper som behandler estetiske uttrykk, en tendens til å basere sin kunnskap på opplevelsene og vurderingene til den skolerte/trente mottakeren. I en rekke sammenhenger er fokuset på den skolerte både fullt berettiget og meget interessant, men ved et ensidig fokus står man i fare for å gå glipp av mye spennende. Opplevelsene av spatiale trekk i reprodusert lyd beveger seg i spenningsfeltet mellom det bevisste og det ubevisste. Skillet mellom den skolerte og den uskolerte lytteren er at førstnevnte er bedre i stand til å flytte klanglige inntrykk fra det ubevisste sanseapparatet til den bevisste erkjennelsen. Målet med oppgaven har vært å få litt mer innblikk i hvordan den uskolerte lytteren evner å erkjenne de stimuli man kan formode hørselens basale mekanismer fanger opp. Et grunnleggende trekk ved populærmusikken er at den favner bredt i befolkningen. Det er min mening at vitenskapen bør forsøke å tegne et bilde som er like bredt.

Det er viktig å merke seg at det er de samme premissene som ligger til grunn for opplevelsene til både den uskolerte og den skolerte lytteren. Utgangspunktet for opplevelsen av spatiale trekk i reprodusert musikk må nødvendigvis ligge i de rammebetingelsene som akustikken og psykoakustikken legger. Etter innledende deler om problemstilling og metode var det derfor naturlig å vie kapittel 2 til de premissene som ligger til grunn for opplevelsen av spatiale trekk. Fokus i kapittelet var på de tidlige refleksjonene, men enkelte andre aspekter ble også berørt. Kapittelet startet med et blikk på Griesingers kategorier som beskriver vår opplevelse av romklang generelt. Etter en definisjon av tidlige refleksjoner så vi på akustiske konsekvenser av både den enkelte refleksjon, så vel som flere refleksjoner samlet. Det ble vist hvordan interaksjonen mellom tidlige refleksjoner og direkte lyd kan preges av kamfiltereffekter, og hvordan stående bølger er en utpreget utfordring i de rom der konsumenter ofte lytter til musikk. I fortsettelsen av kapittelet bevegde vi oss over på psykoakustikken og så hvordan tidlige refleksjoner kan påvirke vår opplevelse av spatiale trekk i tilknytning til den enkelte lydkilde, samt opplevelsen av rommet lydkilden befinner seg i. Til slutt så vi at forskningen har påvist en rekke grenseverdier og kriterier som avgjør hvorvidt de tidlige refleksjoner vil ha en betydning for vår opplevelse.

I kapittel 3 bevegde vi oss over til lydstudioet og så hvordan tidlige refleksjoner preger tilblivelsen til en populærmusikkproduksjon. Kapittelet startet med å vise hvordan naturlig skapte tidlige refleksjoner blir fanget opp under innspillingen, og hvordan de potensielt påvirker det endelige resultatet. Det ble vist at de innspilte refleksjonenes karakter bestemmes av innspillingsrommets utforming, så vel som både lydkildens og mikrofonenes posisjon og karakteristikk. Tidlige refleksjoner spiller også en viktig rolle i lydstudioets kontrollrom, som er omgivelsene der lydteknikeren og produsenten foretar sine valg. Vi så at kontrollrommets utforming, samt plassering og form på det tekniske utstyret er avgjørende for optimale lytteforhold. For å unngå uønskede kamfiltereffekter og maskering av innspilte refleksjoner, ønsker man gjerne en lang initial forsinkelse mellom direkte lyd og første tidlige refleksjon i kontrollrommet. Etter å ha sett på de naturlig skapte refleksjonene i lydstudioet, bevegde fokus seg over mot nedmiksing og den kunstige romklangens domene. Det er under nedmiksing musikkproduksjonens spatiale trekk blir fastsatt. Den kunstige romklangen kan bidra sterkt i utformingen av de spatiale trekkene, og det var derfor naturlig å vie en del plass til hvordan den kunstige romklangen skapes. Vi så at tidlige tiders analoge teknikker fremdeles har et fotfeste i denne verden, men at den kunstige romklangen i hovedsak genereres ved digital prosessering. Romklangen genereres med utgangspunkt i to ulike teknikker. De to teknikkene er algoritmebasert romklang og romklang skapt ved bruk av

impulsresponsers og konvulsjon. Vi så at impulsresponsene er fremragende i sin gjenskaping av naturlige omgivelser, men er noe begrenset med hensyn til manipulering av parametere og simulering av ulike posisjoner. Den algoritmebaserte tilnærmingen åpner for større grad av manipulasjon, som i populærmusikkproduksjon gjerne er ønskelig. Det ble lagt vekt på at en gjenskaping av det helt naturlige ikke nødvendigvis er målet for bruken av kunstig romklang i populærmusikken. Til slutt i kapittelet så vi et par eksempler på hvordan lydteknikere og produsenter benytter kunstig romklang i praksis. Vi så at brukeren ikke nødvendigvis skiller like tydelig mellom de tidlige og sene refleksjonene som oppgaven gjør, og vi så at man i populærmusikkproduksjoner gjerne tilfører flere typer romklang til en og samme miks.

Utgangspunktet for kapittel 4 var å belyse hvordan konsumentens lyttesituasjon avviker fra den optimale lyttesituasjonen. Det gjør den på en rekke områder. Kapittelet startet med å se på hva som skjer i det man som lytter beveger seg bort fra den optimale lytteposisjonen sweet spot. Vi så at det i posisjoner utenfor sweet spot vil oppstå en endring av lydbildets spatiale trekk. Endringene fortøner seg ofte som en forringelse, men samtidig er ikke endringene like forutsigbare som man kanskje kunne tro. Vi så at en økning av avstanden til høyttalerne normalt vil innebære at lytterrommets akustikk får større innvirkning på opplevelsen. Konsumentens tendens til å oppholde seg i nærheten av vegger gjør at de tidlige refleksjonenes påvirkning vil være maksimal. For å utdype emnet ytterligere, fortsatte kapittelet med et blikk på hvilke muligheter spatiale trekk i de tre dimensjonene (høyde, bredde og dybde) har til å overleve i posisjoner utenfor sweet spot. Vi så at forskningen har hatt en tendens til å fokusere på breddedimensjonen og at resultater fra forskningen understreker at overlevelsen til spatiale trekk ikke er helt forutsigbar. Avsnittet om dybdedimensjonen ble viet til opplevelsen av avstand, da det var avstand som var fokus for oppgavens lytteforsøk. I avslutningen av kapittelet, tok vi et blikk på noen trekk som preger konsumentenes lyttesituasjon. Høyttalernes plassering, begrenset båndbredde og avspillingsvolum ble trukket fram. Av disse ble høyttalernes plassering viet mest plass. En lang rekke aspekter ble utelatt, aspekter som i seg selv kunne vært mat til mange hovedoppgaver.

Kapittel 5 gav en beskrivelse av de praktiske forberedelsene og gjennomføringen av lytteforsøket som ble gjennomført i forbindelse med denne oppgaven. Etter innspilling i ekkofritt rom, ble musikken mikset i ulike versjoner ved å variere lydstyrken på de tidlige refleksjonene. Nivåforholdet mellom den direkte lyden og de tidlige refleksjonene benyttes av hørselen når den skal bedømme avstanden til en lydkilde. Høyere nivåer på de tidlige refleksjonene vil oppleves som en økning i avstanden, forutsatt at andre faktorer forblir like.



Forsøket var delt inn i 5 oppgaver, hvor hver oppgave besto av 4 versjoner av ett lydeksempel. Forskjellen mellom de 4 versjonene var altså ulikt nivå på de tidlige refleksjonene. Deltakernes oppgave var å vurdere avstanden til instrumentene i hver versjon. I forsøket ble det benyttet en grafisk metode som gjør det mulig for deltakeren å formidle sin opplevelse uten å måtte gå via språket. Deltakerne benyttet et datagenerert grensesnitt med en graderingsskala for å formidle sin opplevelse av avstanden. Fordi opplevd avstand ikke er et kvantifiserbart fenomen, ble tolkningen av graderingsskalaen overlatt til deltakeren selv.

I forberedelsene til forsøket ble det lagt vekt på å få et jevnt avspillingsnivå på de fire versjonene innenfor den enkelte oppgave. Dette ble gjort fordi opplevd lydstyrke er en meget dominant faktor i hørselens avstandsvurdering. Dessverre ble det i overføringen til datagrensesnittet gjort en feil. Feilen ble ikke vurdert til å ha noen betydning for forsøket som helhet, men resulterte i at to versjoner ble utelatt fra analysen av deltakernes besvarelser.

Analysen av deltakernes besvarelser utgjør oppgavens sjettede kapittel. Sett i lys av undertegnedes ønske om å tilføre musikkvitenskapen ny kunnskap er kapittel 6 det desidert viktigste. Det må bli opp til andre å vurdere om jeg faktisk har generert ny kunnskap, eller om jeg bare har bidratt til forvirring. Jeg skal peke på et par ting jeg mener vi har lært litt om, men det meste av fokuset i avslutningen av denne oppgaven skal jeg vie til spørsmålet om hvordan vi kan lære mer.

Har vi så lært noe om konsumenters opplevelse av spatiale trekk i løpet av forsøket og analysen? Jeg skal vokte meg for å si at vi har lært noe helt sikkert, men jeg vil driste meg til å si at vi aner konturene av noe. Det overordnede inntrykket av forsøket er at besvarelsene spriker, men at trendene peker i retning av hypotesene som ble framsatt. Skulle jeg forsøke å trekke ut essensen, blir det at konsumenter evner å oppfatte nyanser av spatiale trekk i reprodusert populærmusikk, men de må arbeide for å klare det. Et meget relevant spørsmål blir da om konsumentene noen gang forsøker å oppfatte spatiale trekk i det hele tatt. Oppgaven har ikke forsøkt å svare på dette spørsmålet, men det er en del som tyder på at konsumentene sjelden gjør det. En bevisst opplevelse av spatiale trekk handler om at lytteren må rette sitt fokus mot musikken. Undersøkelser har vist at konsumentenes musikkbruk i stadig større grad preges av å være et bakteppe, mens lytteren har sitt fokus helt andre steder (Foss, 2006). Med utgangspunkt i kunnskapen om konsumenters lytting til musikk, kan man kritisere oppgavens problemstilling for å ha feil fokus. I stedet for å fokusere på hva konsumenter klarer å gripe med bevisstheten, burde man kanskje heller ha fokusert på hvilke spatiale trekk konsumenten ubevisst opplever. Her er undertegnede helt enig (med seg selv:-). Det er her det virkelig interessante ligger. Problemet er at det synes å være en lang vei å gå før

vi kommer dit. Denne oppgaven kan således leses som et lite steg på veien til den ubevisste opplevelsen, og de basale mekanismene som ubønhørlig må prege vår opplevelse av musikkens spatiale trekk.

I forbindelse med spørsmålet om hvordan vi kan lære mer er det naturlig å se på forbedringspotensialet til oppgavens lytteforsøk. Med andre ord, kunne vi lært mer dersom vi hadde gjort ting annerledes? Svaret er definitivt ja. Nedenfor følger en liten liste med områder der undertegnede i etterkant ser muligheter for forbedring.

1. I tillegg til konsumenter som deltakere burde forsøket vært kjørt på en gruppe skolerte lyttere. Ved å sammenstille svarene fra skolerte og uskolerte, kunne man utelukket noen potensielle feilkilder. Samtidig kunne man plassert konsumentens opplevelse i et tydeligere lys.
2. Nivåene på de tidlige refleksjonene var som kjent tilfeldig valgte. I forkant av forsøket kunne man ved bruk av ABX-forsøk avdekket grenseverdier for hva konsumentene evner å skille ved ulike posisjoner. Resultatene fra et slikt ABX-forsøk kunne vært benyttet i nivåsettingen av de tidlige refleksjonene. En slik framgangsmåte ville økt forsøkets presisjonsnivå, og dermed gitt større relevans og kunnskap.
3. Antallet deltakere i forsøket kunne med fordel vært langt høyere.
4. Nivåjusteringen burde vært enda grundigere, og kunne med fordel vært utført i samarbeid med flere kvalifiserte lyttere.
5. Man må anta at et bedre tilpasset grensesnitt hadde gitt bedre resultater, men hvordan man får til dette i opplevelsen av avstand er vanskelig å si.
6. Forsøket kunne vært avholdt i sertifisert lokale.

Noe av forbedringspotensialet var undertegnede fullstendig klar over på forhånd, men den store utfordringen er at forsøk av denne art er meget tidkrevende. I mangelen av tilgang til et bedre egnet forsøkslokale ble forsøket avholdt i undertegnedes egen stue. Den primære oppgaven til denne stuen er å fungere som stue, ikke forsøkslokale. Dette la klare begrensninger på antallet deltakere i forsøket, og også forberedelsene til forsøket.

Er det så noen mulighet for at vitenskapen ønsker å lære mer om konsumenters opplevelse av spatiale trekk i musikken? Det er god grunn til å håpe det. Det er et stort lyspunkt at et kjent navn som Begault har tatt tak i noe av samme problemstilling (Begault, 2006). Følgende er en liten avsporing, men Begault peker blant annet på et meget interessant aspekt ved opplevelsen av spatiale trekk. Vi mennesker er i stadig bevegelse, også når vi lytter til musikk. Man må anta at vårt inntrykk av musikkens spatiale trekk dannes på

bakgrunn av erfaring fra flere posisjoner. Det hele synes meget innlysende, men det er en utfordring for vitenskapen å skulle påvise det. Grafiske metoder synes å være et velegnet verktøy i arbeidet med å avdekke opplevelser av spatiale trekk, og må antas å kunne benyttes også til dette. Vi har sett at oppgavens grafiske metode er hentet fra et fagfelt som retter seg inn mot lyd kvalitet. Fagfeltet er tett forbundet med undertegnedes innfallsvinkel. Det er gledelig å se at flere og flere publikasjoner fra The Audio Engineering Society retter seg inn mot ikke-optimale lyttesituasjoner. Vi får håpe det smitter<sup>1</sup>.

---

<sup>1</sup> Slutt

## Litteraturliste

- Bech, S. 1999, "Methods for subjective evaluation of spatial characteristics of sound", *The Proceedings of The Audio Engineering Society 24th International Conference: Multichannel Audio The New Reality* pp171-183.
- Bech, S. 1998, "Spatial aspects of reproduced sound in small rooms", *J. Acoustical Society of America*, 103, January 1998.
- Begault, D.R., McClain, B.U., Anderson, M.R. 2004, *Early Reflection Thresholds for Anechoic and Reverberant Stimuli within a 3-D Sound Display*, Human Factors Research and Technology, NASA, USA, tilgjengelig fra: <[http://human-factors.arc.nasa.gov/publibrary/20050322115330\\_Begault\\_ICA-finalREVISED.pdf](http://human-factors.arc.nasa.gov/publibrary/20050322115330_Begault_ICA-finalREVISED.pdf)>, [mai 2005]
- Begault, D.R. 2006, "Preference versus Reference: Listeners as participants in sound reproduction", *Spatial Audio and Sensory Evaluation Techniques*, Guildford, UK, 2006 april 6-7, tilgjengelig fra: <<http://www.surrey.ac.uk/soundrec/ias/papers/Begault.pdf>>, [august 2006]
- Beranek, L., Griesinger, D. 2004, "Listening to concert halls", *Acoustical Society workshop*, tilgjengelig fra: <<http://www.std.com/~griesngr/asa04.ppt>>, [mai 2005]
- Birkeland, T. 1998, *Personlig intervju*,
- "Bokmålsordboka" 2006, tilgjengelig fra <<http://www.dokpro.uio.no/ordboksoek.html>>, [januar 2006]
- Bozzoli, F., Farina, A. 2003, "Directivity balloons of real and artificial mouth simulators for measurement of the Speech Transmission Index", *Audio Engineering Society 115th Convention*, Convention paper, tilgjengelig fra: <<http://pcfarina.eng.unipr.it/Public/Papers/186-AES115.PDF>>, [januar 2006]
- Bregman, A. 1990, *Auditory scene analysis*, Cambridge, Mass, MIT Press.
- Cabrera, D. 1999, "Psysound: A computer program for psychoacoustical analysis", *Proceedings of the Australian Acoustical Society Conference*, Melbourne, 24-26 November 1999, pp 47-54.
- "Caplex" 2006, tilgjengelig fra <<http://www.caplex.no>>, [tilgang: januar 2006]
- Davis, D., Davis, C. 1997, *Sound system engineering*, 2. utgave, Boston : Focal Press.
- Devore, S., Shinn-Cunningham, B. 2003, "Perceptual consequences of including reverberation in spatial auditory displays", *Proceedings of the 2003 International Conference on Auditory Display*, Boston, Ma, USA, 6-9 July 2003, tilgjengelig fra: <[http://cns-web.bu.edu/~shinn/pages/pdf/ICAD\\_03\\_Devore.pdf](http://cns-web.bu.edu/~shinn/pages/pdf/ICAD_03_Devore.pdf)>, [mai 2005]

- Engine Ears 2002, "Debattforum", tilgjengelig fra: <<http://www.engine-ears.no/cgi-bin/YaBB/YaBB.pl?board=Studio;action=display;num=1036143878>>, [august 2006]
- Everest, F. Alton 2001, *The Master handbook of acoustics*, 4. utgave, New York : McGraw-Hill.
- Foss, A.S. 2006, *Full musikk, null opplevelse?*, Forskning.no, tilgjengelig fra: <<http://www.forskning.no/Artikler/2006/februar/1139233318.3>>, [september 2006]
- Gardner, W., Griesinger, D. 1994, "Reverberation Level Matching Experiments", *Acoustical Society of America, Proceedings of the Sabine Memorial Conference*, MIT, June 1994, tilgjengelig fra: <<http://www.std.com/~griesngr/gardner.doc>>, [mai 2005]
- Gardner, W. 1997, *Privat kommunikasjon*
- Gardner, W. 1996, "Reverberation algorithms", Utkast til publikasjon i *Applications of signal processing to audio and acoustics*, Kluwer Academic Press (formidlet privat).
- Gardner, W. 1992, *The virtual acoustic room*, Master-thesis, Massachusetts Institute of Technology.
- Genelec 2006, "FAQ", tilgjengelig fra: <<http://www.genelec.com/support/faq/faq21.php>>, [januar 2006]
- Geoffrey, G.M. 2001, *A hybrid model for simulating diffused first reflections in two-dimensional acoustic environments*, PhD-avhandling, Faculty of Music, McGill University, tilgjengelig fra: <<http://www.tonmeister.ca/research/pubs/martin01.pdf>>, [januar 2006]
- Griesinger, D. 2005a, *Recent experiences with electronic acoustic enhancement in concert halls and opera houses*, tilgjengelig fra: <<http://www.std.com/~griesngr/icsv.pdf>>, [mai 2005]
- Griesinger, D. 2005b, *Recent experiences with electronic acoustic enhancement in concert halls and opera houses - lecture slides*, tilgjengelig fra: <<http://www.std.com/~griesngr/iscv.pdf>>, [mai 2005]
- Griesinger, D. 2000, "The Theory and Practice of Perceptual Modeling - How to use Electronic Reverberation to Add Depth and Envelopment Without Reducing Clarity", *Preprint from Tonmeister conference in Hannover, 2000*, tilgjengelig fra: <<http://www.std.com/~griesngr/threedpm.pdf>>, [mai 2005]
- Griesinger, D. 1999, "Objective Measures of Spaciousness and Envelopment", *Audio Engineering Society, Proceedings of the 16th international conference on Spatial Sound Reproduction*, pp 27-41, tilgjengelig fra: <<http://www.std.com/~griesngr/objmeas.pdf>>, [mai 2005]
- Griesinger, D. 1998a, "Multichannel sound systems and their interaction with the room", *Audio Engineering Society, Proceedings of the 15th International Conference of the*

- AES on small room acoustics*, pp 159-173., tilgjengelig fra:  
<<http://www.std.com/~griesngr/multichan.pdf>>, [mai 2005]
- Griesinger, D. 1998b, "General Overview of Spatial Impression, Envelopment, Localization, and Externalization", *Audio Engineering Society, Proceedings of the 15th International Conference of the AES on small room acoustics*, pp 136-149, tilgjengelig fra: <<http://www.std.com/~griesngr/overvw1.pdf>>, [mai 2005]
- Griesinger, D. 1997, "The psychoacoustics of apparent source width, spaciousness & envelopment in performance spaces", *Acta Acustica* Vol. 83 (1997) 721-731, tilgjengelig fra: <<http://www.std.com/~griesngr/SPAC7A.DOC>>, [mai 2005]
- Griesinger, D. 1995, "How loud is my reverberation", *Audio Engineering Society 98th Convention*, Convention paper, tilgjengelig fra:  
<[http://www.std.com/~griesngr/how\\_loud2.doc](http://www.std.com/~griesngr/how_loud2.doc)>, [mai 2005]
- Griesinger, D. 1989, "Practical Processors and Programs for Digital Reverberation", *Audio Engineering Society, Proceedings of the 7th International Conference*, Toronto May 1989, tilgjengelig fra: <<http://www.std.com/~griesngr/prctpro.pdf>>, [mai 2005]
- Guastavino, C., Katz, B.F.G. 2004, *Perceptual evaluation of multi-dimensional spatial audio reproduction*, Acoustical Society of America, tilgjengelig fra:  
<<http://www.gslis.mcgill.ca/guastavino/publis/GuastavinoKatz2004JASA.pdf>>, [mai 2005]
- Halmrast, T. 1999, "Musikal timbre combfilter-coloration from reflections", *Proceedings of the 2nd COST G-6 Workshop on Digital Audio Effects (DAFx 99)*, NTNU, Trondheim, December 9-11, 1999.
- Hameed, S., Pakarinen, J., Valde, K., Pulkki V. 2004, "Psychoacoustic Cues in Room Size Perception", *Audio Engineering Society 116th Convention*, Convention paper 6084, tilgjengelig fra:  
<<http://www.acoustics.hut.fi/research/cat/psychoac/papers/hameedaes116.pdf>>, [desember 2005]
- Hammer, Ø. 1997, *Digital lydbehandling*, Norsk nettverk for Teknologi, Akustikk og Musikk.
- Hammer, Ø. 1995, *Akustikk for musikere*, Norsk nettverk for Teknologi, Akustikk og Musikk.
- Kapralos, B. 2003, *Auditory Perception and Virtual Environments*, PhD-avhandling, York University, North York, Ontario, Canada, tilgjengelig fra:  
<<http://www.cse.yorku.ca/~billk/billkPhdPrelim.pdf>>, [januar 2006]
- LaCarrubba, M. 1999, *Making Loudspeakers And Control Rooms That Make Music "Sound Good"* tilgjengelig fra:  
<[http://www.moultonlabs.com/more/making\\_music\\_sound\\_good/](http://www.moultonlabs.com/more/making_music_sound_good/)>., [august 2006]
- Lexicon 2001, *Lexicon 960L Digital Effects System - Owners manual*, rev 2, tilgjengelig fra:  
<[http://www.lexiconpro.com/Product\\_Downloads/960LLARC2/960L\\_Owners\\_Manual\\_Rev2.pdf](http://www.lexiconpro.com/Product_Downloads/960LLARC2/960L_Owners_Manual_Rev2.pdf)>, [tilgang mai 2005]

- Lexicon 2000, *This is the future*, Brosjyre for Lexicon 960L, Lexicon, tilgjengelig fra: <<http://www.lexiconpro.com>>, [april 2005]
- Lexicon 1993, *Lexicon 480L Digital Effects System - Owners manual*, rev 0, tilgjengelig fra: <[http://www.lexiconpro.com/Product\\_Downloads/480L\\_V4\\_Owners\\_Manual\\_Rev0.pdf](http://www.lexiconpro.com/Product_Downloads/480L_V4_Owners_Manual_Rev0.pdf)>, [tilgang september 2006]
- McAdams, S., Bigand, E. 1993, *Thinking in sound : the cognitive psychology of human audition*, Oxford : Clarendon Press.
- Moore, B.C.J., Glasberg, B.R., Baer, T. 1997, "A model for the prediction of thresholds, loudness and partial loudness", *J. Audio Eng. Soc.* Vol 45, No. 4 1997, tilgjengelig fra: <<http://www.aes.org/e-lib/browse.cfm?elib=10272>>, [juni 2006]
- Moore, F. Richard 1990, *Elements of computer music*, Englewood Cliffs, N.J. : Prentice Hall.
- Moorer, James A. 1979, "About this reverberation business", *Computer Music Journal*, Volume 3, Number 2, 1979, pp13-28.
- Roads, Curtis 1996, *The computer music tutorial*, Cambridge, Mass. : MIT Press.
- Rossing, T.D., Moore, R.F., Wheeler, P.A. 2002, *The science of sound*, 3. utgave, San Francisco : Addison Wesley.
- Rumsey, Francis 2006, *Privat kommunikasjon*
- Rumsey, F., Lee, H-K. 2004, "Elicitation and grading of subjective attributes of 2-channel phantom images", *Audio Engineering Society 116th Convention*, Convention paper 6142, tilgjengelig fra: <<http://epubs.surrey.ac.uk/recording/15>>, [januar 2006]
- Rumsey, F., Mason, R., Brookes, T. 2003, "Creation and verification of a controlled experimental stimulus for investigating selected perceived spatial attributes", *Audio Engineering Society 114th Convention*, Convention paper, tilgjengelig fra: <<http://epubs.surrey.ac.uk/recording/11>>, [januar 2006]
- Rumsey, F., Ford, N., Nind, T. 2003a, "Creating a universal graphical assessment language for describing and evaluating spatial attributes of reproduced audio events", *Audio Engineering Society 115th Convention*, Convention paper 5907, tilgjengelig fra: <<http://epubs.surrey.ac.uk/recording/13>>, [januar 2006]
- Rumsey, F., Ford, N., Nind, T. 2003b, "Evaluating spatial attributes of reproduced audio events using a graphical assessment language - understanding differences in listener depictions", *Audio Engineering Society 24th International Conference on Multichannel Audio*, tilgjengelig fra: <<http://epubs.surrey.ac.uk/recording/25>>, [januar 2006]
- Rumsey, F., Mason, R. 2002, "A comparison of objective measurements for predicting selective subjective spatial attributes", *Audio Engineering Society 112th Convention*,

- Convention paper 5547, tilgjengelig fra: <<http://epubs.surrey.ac.uk/recording/10>>, [januar 2006]
- Rumsey, F., Ford, N., Nind, T. 2002, "Subjective evaluation of perceived spatial differences in car audio systems using a graphical assessment language", *Audio Engineering Society 112th Convention*, Convention paper 5547, tilgjengelig fra: <<http://epubs.surrey.ac.uk/recording/30>>, [januar 2006]
- Rumsey, F., Neher, T., Brookes, T. 2002, "Training of listeners for the evaluation of spatial sound reproduction", *Audio Engineering Society 112th Convention*, Convention paper 5584, tilgjengelig fra: <<http://epubs.surrey.ac.uk/recording/28>>, [januar 2006]
- Rumsey, F., Bruyn, B. de, Ford, N. 2001, "Graphical elicitation techniques for subjective assessment of the spatial attributes of loudspeaker reproduction – a pilot investigation", *Audio Engineering Society 110th Convention*, Convention paper 5388, tilgjengelig fra: <<http://epubs.surrey.ac.uk/recording/39>>, [januar 2006]
- Rumsey, Francis 2001, *Spatial audio*, Oxford : Focal Press.
- Sheaffer, J., Elyashiv, E. 2005, *Implementation of Impulse Response Measurement Techniques*, Waves Audio Ltd., and Jordan Valley College, tilgjengelig fra: <<http://www.acoustics.net/objects/pdf/IR-paper.pdf>>, [mai 2006]
- Shinn-Cunningham, B., Ihlefeld, A. 2004, "Selective and divided attention: extracting information from simultaneous sound sources", *Proceedings of IACD 04-Tenth Meeting of the International Conference on Auditory Display*, Sydney, Australia, July 6-9, 2004, tilgjengelig fra: <[http://www.icad.org/websiteV2.0/Conferences/ICAD2004/papers/shinn-cunningham\\_etal.pdf](http://www.icad.org/websiteV2.0/Conferences/ICAD2004/papers/shinn-cunningham_etal.pdf)>, [mai 2005]
- Shinn-Cunningham, Barbara 2003, "Acoustics and perception of sound in everyday environments", *Proceedings of the 3rd Int. Workshop on Spatial Media*, Aizu-Wakamatsu, Japan, Mar. 6-7, 2003, tilgjengelig fra: <<http://cns-web.bu.edu/~shinn/pages/pdf/Aizu.pdf>>, [mai 2005]
- TCElectronic 2005, *System 6000 - Algorithms*, rev 3.5, tilgjengelig fra: <[http://www.tcelectronic.com/media/4%20Sys6\\_Algos\\_350\\_1612.pdf](http://www.tcelectronic.com/media/4%20Sys6_Algos_350_1612.pdf)>, [tilgang september 2006]
- Thagaard, Tove 2004, *Systematikk og innlevelse - en innføring i kvalitativ metode*, 2. utgave, Fagbokforlaget.
- Traunmüller, H., Eriksson, A. 1995, *The frequency range of the voice fundamental in the speech of male and female adults*, Institutionen för lingvistik, Stockholms universitet, tilgjengelig fra: <[http://www.ling.su.se/staff/hartmut/f0\\_m&f.pdf](http://www.ling.su.se/staff/hartmut/f0_m&f.pdf)>, [februar 2006]
- Vaage, Odd Frank 2005, *Norsk mediebarometer 2005*, Statistisk sentralbyrå, Oslo, tilgjengelig fra: <<http://www.ssb.no/emner/07/02/30/medie/sa78/sa-78.pdf>>, [mai 2006]



Wave Arts 2006, *Panorama – 3D virtual acoustics*, Wave Arts Inc, tilgjengelig via:  
<<http://www.wavearts.com/Panorama5.html>>, [september 2006]

Walker, R. 1995, *Controlled image design: The management of stereophonic image quality*,  
Research & Development Department, Technical Resources Division, BRITISH  
BROADCASTING CORPORATION, tilgjengelig fra:  
<<http://www.bbc.co.uk/rd/pubs/reports/1995-04.pdf>>, [desember 2005]

Warner, Timothy 2003, *Pop music, technology and creativity : Trevor Horn and the digital  
revolution*, Aldershot : Ashgate.

White, Paul 2001a, *Advanced reverberation: Part 1*, tilgjengelig fra:  
<<http://www.soundonsound.com/sos/oct01/articles/advancedreverb1.asp>>, [Publisert i  
Sound on Sound oktober 2001]

White, Paul 2001b, *Advanced reverberation: Part 2*, tilgjengelig fra:  
<<http://www.soundonsound.com/sos/nov01/articles/advancedreverb2.asp>>, [Publisert i  
Sound on Sound november 2001]

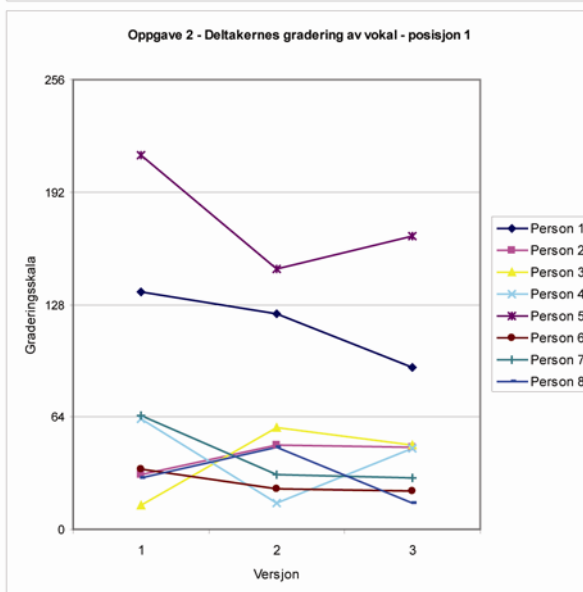
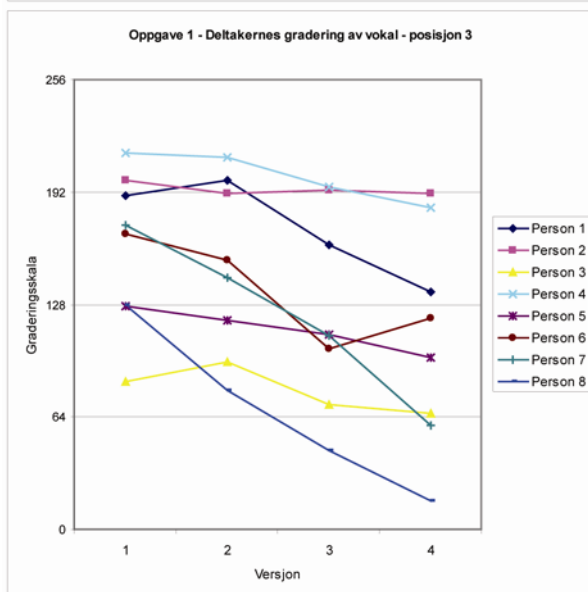
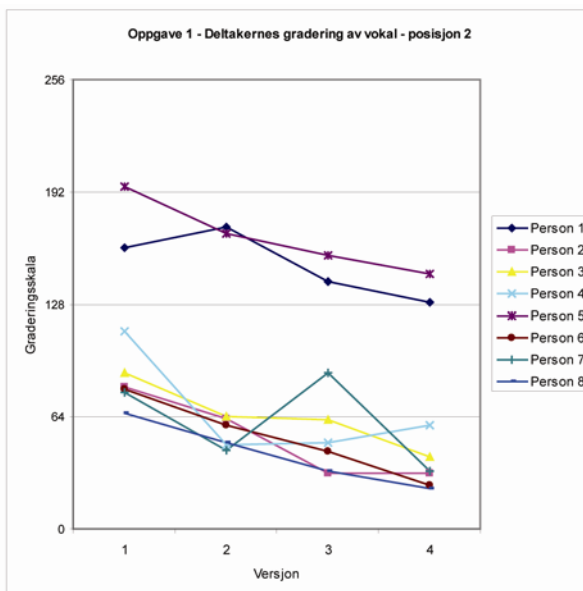
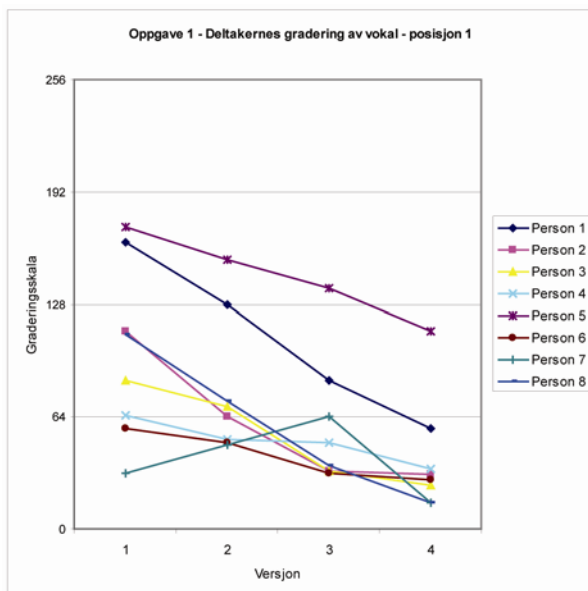
”Wikipedia” 2006, tilgjengelig fra: <<http://no.wikipedia.org>>, [september 2006]

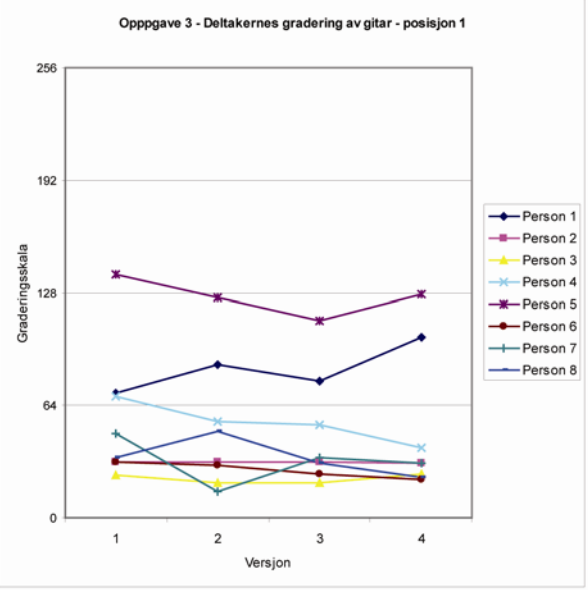
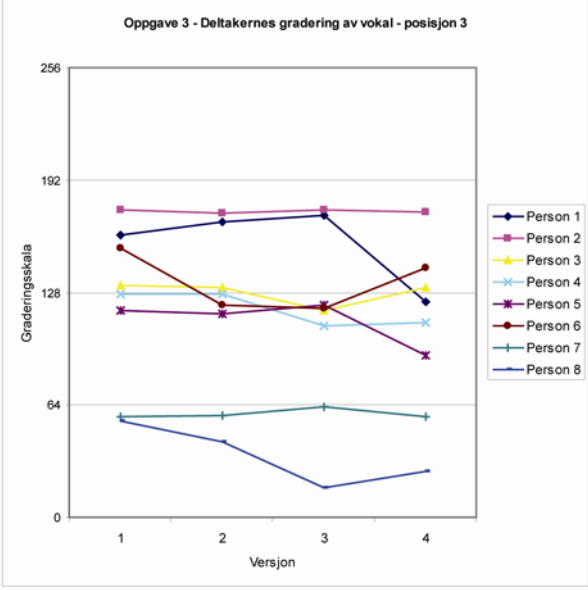
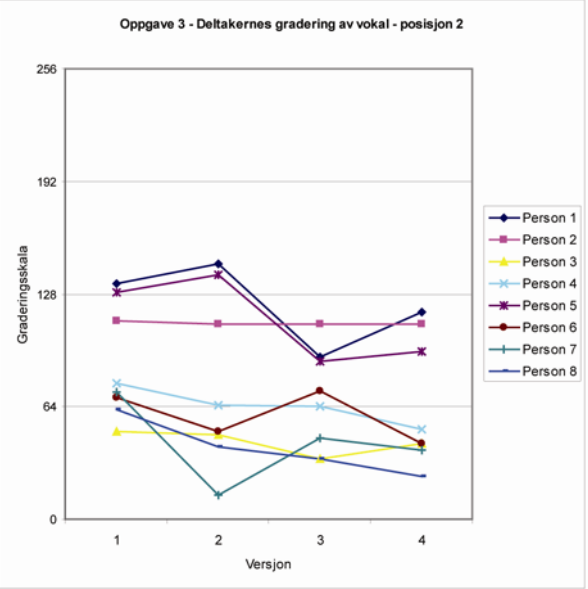
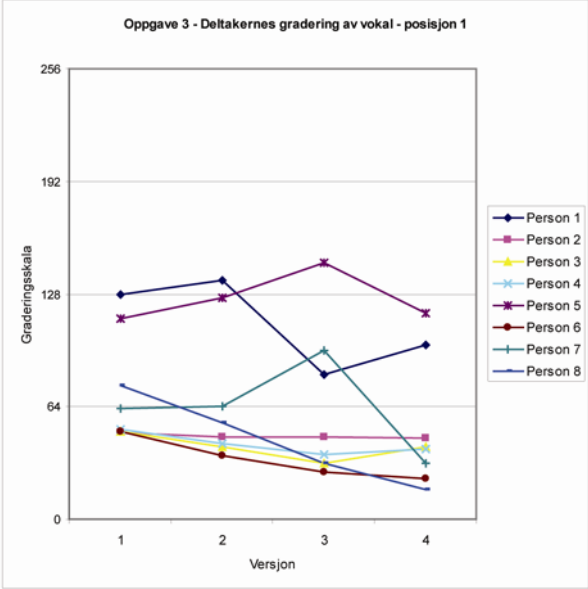
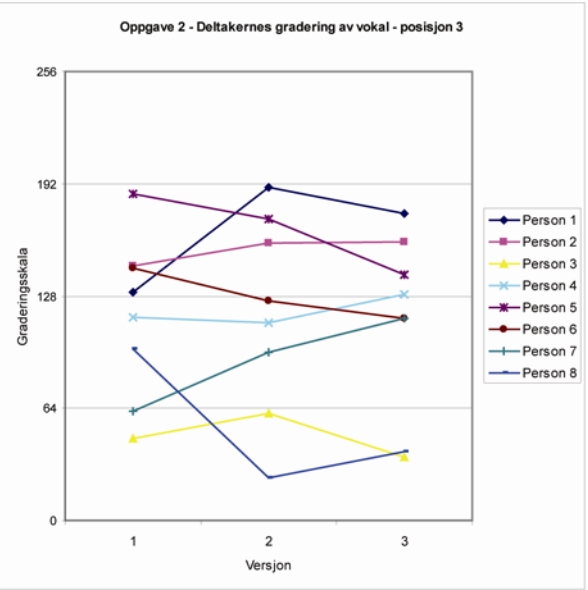
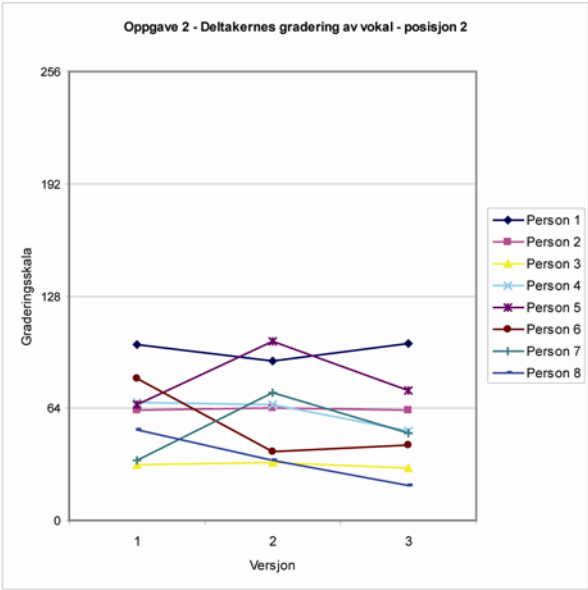
### **Programvare**

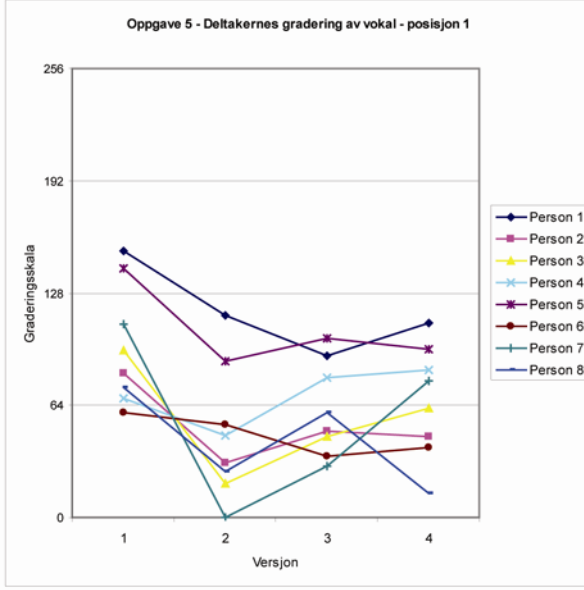
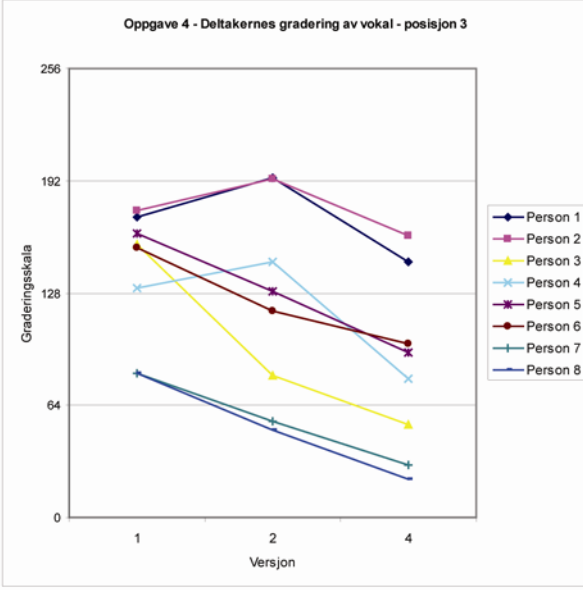
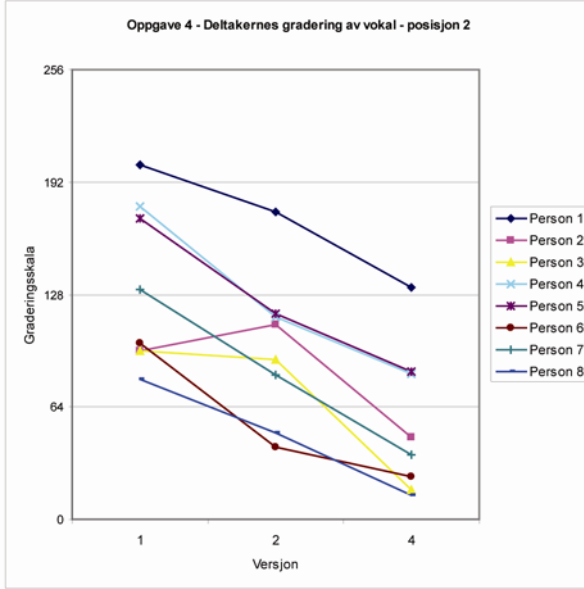
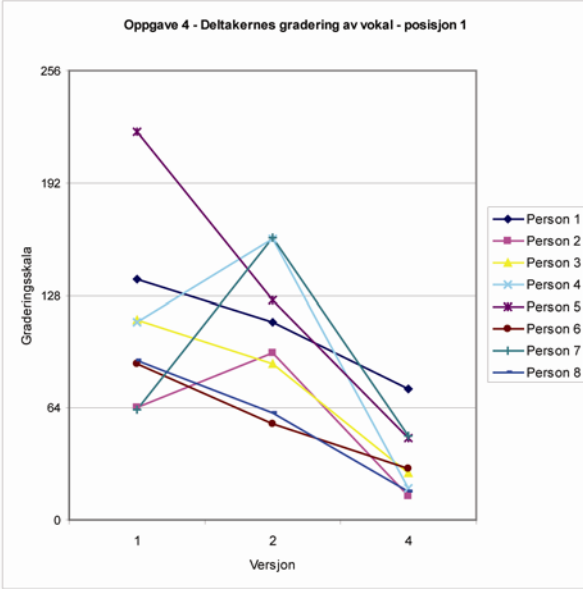
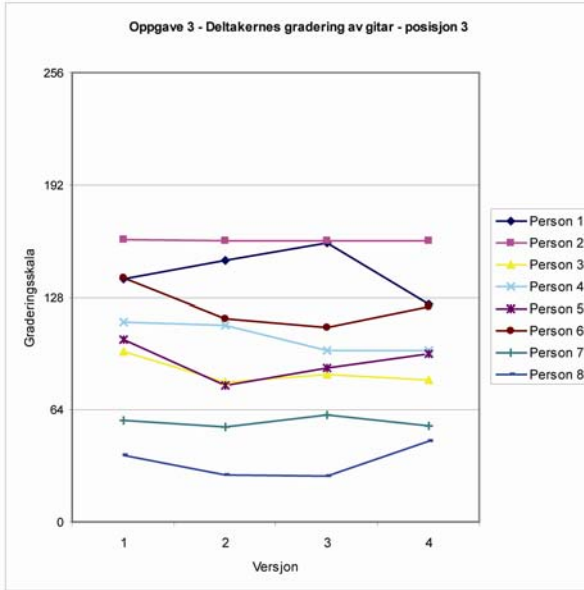
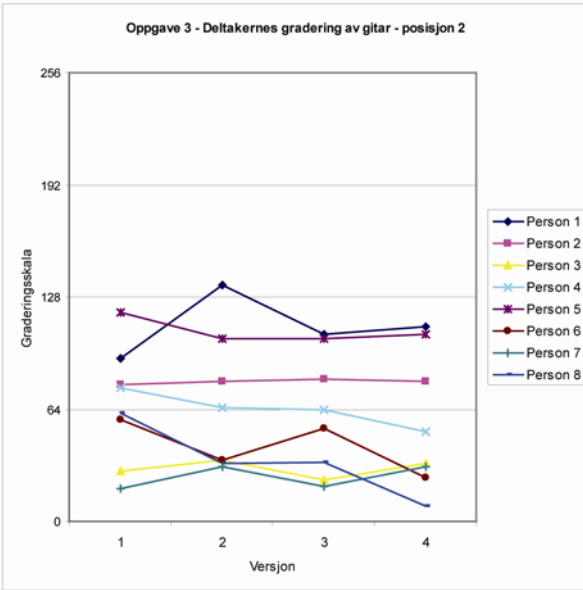
”PCABX” 2005, *Dataprogram*, <<http://www.pcabx.com>>, [juni 2005]

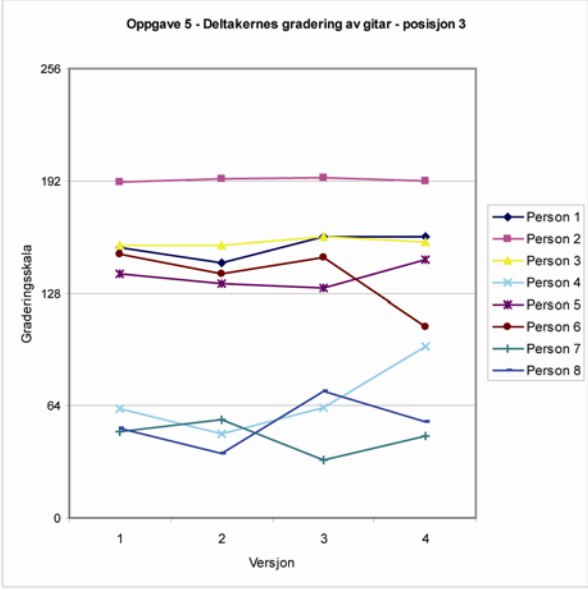
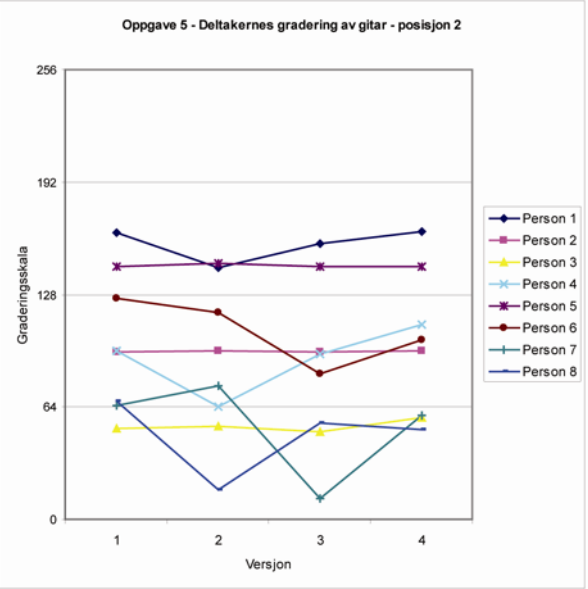
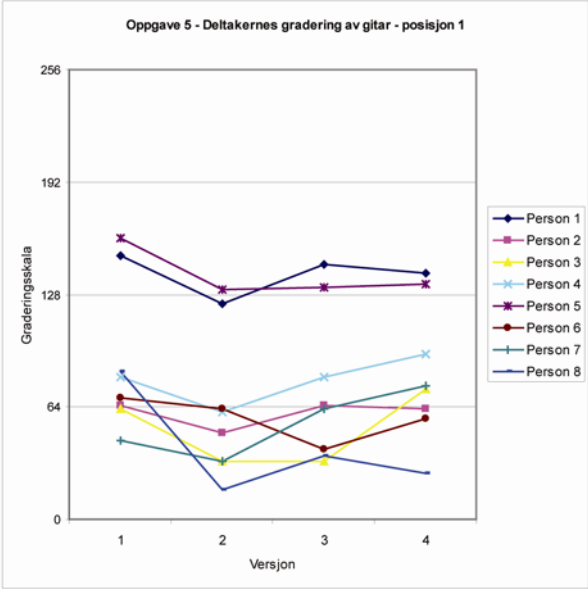
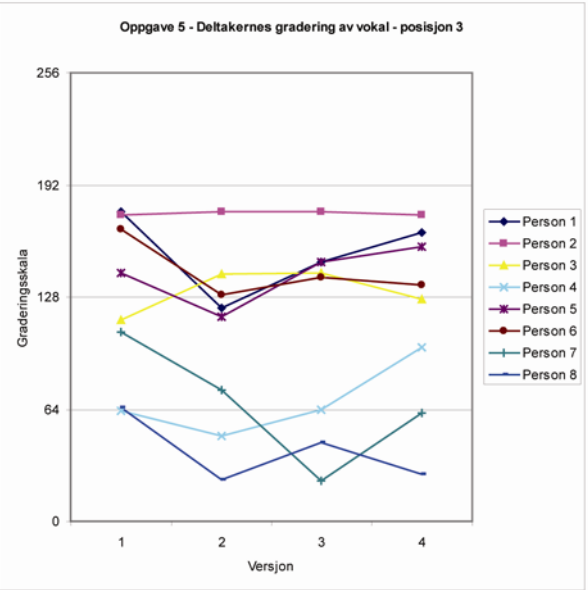
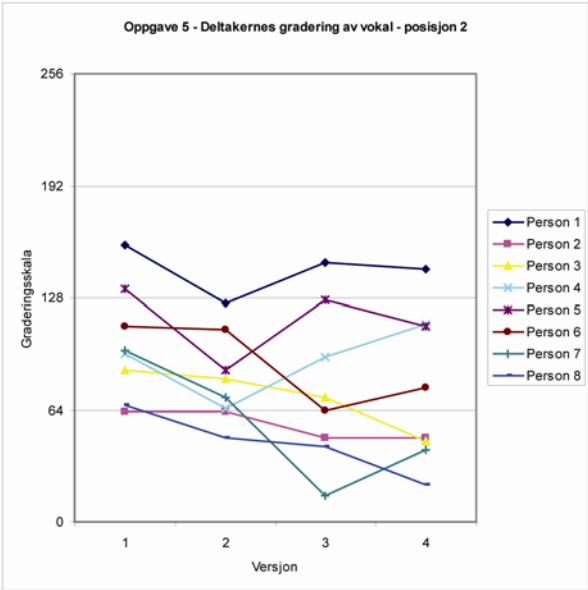
# Appendiks A

Appendiks 1 er en samling av alle oppgavebesvarelsene i lytteforsøket. Diagrammene er fordelt på oppgave, posisjon og instrument.









## Appendiks B

Vedlagt med oppgaven er en data-cd som inneholder følgende filer:

1. Lytteforsøkets 15 oppgaver slik de ble presentert for deltakerne. For å kunne spille av filene, må en ha programmet Adobe Flash Player installert. (For nedlasting av Adobe Flash Player gå til: <http://www.adobe.com/products/flashplayer/>)
2. En tekstfil "Les meg.txt" som inneholder teknisk informasjon for avspilling av innholdet på cd'en.
3. En mappe med kildekoden til flash-applikasjonen, php-scriptet som mottar svarene, og tabellstruktur i MySQL.