



UPPLEVD LJUDKVALITET MED TEKNISK KOPPLING

En undersökning i kvalitetsförlust &
människans uppfattningsförmåga

Examensarbete inom huvudområdet
Datalogi/Datavetenskap/Medier, estetik och
berättande
Grundnivå 30 högskolepoäng
Vårtermin 2012

Marcus Nykvist

Handledare: Anders Sjölin
Examinator: Toivo Burlin

Sammanfattning

Arbetet innefattar en bakgrund i de byggstenar som finns i en digital ljudfil och hur de påverkar ljudet. De problem som undersökningen menar att besvara är när upplever en människa kvalitetsförlust och vilken del eller aspekt av ljudet är det viktigaste för upplevelsen. Även huruvida ljud av olika ursprung som syntbaserade eller inspelade ljud påverkas i olika grad av kvalitetsförlust. De försökspersoner som använts vid undersökningen har olika bakgrund, ena gruppen har erfarenhet av ljuddesign eller Hifi medan den andra inte har speciella preferenser beträffande ljud. Undersökningen genomfördes med en kvantitativ undersökning där försökspersonernas uppfattningsförmåga testades och sedan sammanställdes. Resultatet gav en medelpunkt för kvalitet som kan användas för ljuddesign för mobila plattformar. Resultatet används vid ett företag i Skövde där det implementeras som en del av ljuddesign för mobilspel.

Nyckelord: Ljudkvalitet, Mobilplattform, Ljuddesign.

Innehållsförteckning

1	Introduktion	1
2	Bakgrund	2
2.1	Termer	3
2.1.1	Samplingsfrekvens	3
2.1.2	Bitdjup & kvantisering	3
2.1.3	Format.....	4
2.2	Tidigare forskning	5
2.2.1	Vetenskaplig litteratur	6
2.2.2	Övrig litteratur.....	7
3	Problemformulering	8
3.1	Metodbeskrivning	9
3.2	Varianter av ljud	9
3.3	Försökspersoner	10
3.4	Utrustning	11
3.5	Undersökningen	12
3.6	Utvärdering	12
4	Genomförande	14
4.1	Praktiskt genomförande	15
4.2	Förstudie	17
5	Analys	19
5.1	Mönster	20
5.2	Statistik	21
5.2.1	Ljud: Inspelad explosion	22
5.2.2	Ljud: Syntbaserad explosion	23
5.2.3	Ljud: Inspelat effektljud.....	24
5.2.4	Ljud: Syntbaserat effektljud	25
5.2.5	Ljud: Inspelat tal	26
5.3	Summering	27
5.3.1	Jämförelse av ljudtyper	27
5.3.2	En medelpunkt	29
6	Slutsats	30
6.1	Resultatsammanfattning	30
6.2	Diskussion	31
6.3	Framtida arbete	32

1 Introduktion

Detta arbete fokuserar på ljudkvalitet och vad det är som påverkar upplevd förlust av kvalitet. Det innehåller delar av både den tekniska bakgrunden som behövs för att förstå vad som händer rent tekniskt samt en ytlig förklaring i hur människor faktiskt upplever kvalitetsförlust och i vilket skede, och framförallt vilken beståndsdel av ett ljud påverkar den upplevda kvalitén mest. Arbetet kommer att handla om upplevd kvalitet och inte teknisk kvalitet, förutom att försöka hitta en kritisk nivå när en person upplever att förlust har inträffat och sedan notera detta för framtida användningsområden. Arbetet är menat att kunna användas till framförallt mobilspel, där det är vitalt att hitta en balans mellan ljudkvalitet och filstorlek för att få ut så mycket som möjligt på en plattform med restriktioner.

Digitala ljud är en konvertering av analoga signaler som plockas upp av någon typ av mikrofon. Kvalitén på dessa signaler beror på hur det samplas, mer specifikt hur frekvent det samplas. Sampling kan ses som en typ av stickprov på en analog signal där man gör ett test på hur signalen låter just vid det tillfället, och sedan igen vid ett senare tillfälle och så vidare. På detta sätt så kan man se skillnader på de två punkterna och hur ljudet förändras. Detta är dock en grov förenkling av ett system som arbetar väldigt snabbt för att få fram trovärdiga ljud. Senare i bakgrunden så förklaras sampling på ett mer ingående med exempelbilder för att förenkla förståelsen för hur det fungerar.

Problemet som undersökts är när en människa upplever kvalitetsförlust i en digital ljudfil och vilken teknisk aspekt av ljudet som påverkas mest. De två aspekter som undersöks är bitdjup och samplingsfrekvens och hur de påverkar ljudet vid en minskning av deras värden. Även om ljud av olika ursprung påverkas på olika sätt och på så sätt är lämpade för olika typer av kvalitet. Med ursprung så menas det sättet de skapats på, till exempel genererat med en mjukvarusynt eller inspelat med en mikrofon. Detta inkluderades då det är av intresse att se huruvida olika typer av ljud påverkas i olika grad.

De ljud som användes i undersökningen skapades med hjälp av både en mjukvarusynt samt inspelat material som spelades in med hjälp av en portabel inspelare samt en separat dynamisk mikrofon. Fem olika ljud som representerar de olika kategorierna av ljud användes, varje ljud har 12 steg i kvalitet. Varje steg i kvalitet bedömdes av försökspersonerna med en femgradig skala om hur mycket ljudkvalitén hade försämrats från referensljudet som var det ljud som hade högsta möjliga ljudkvalitén för undersökningen.

2 Bakgrund

Detta arbete är ett resultat av intresse samt behov av ny kunskap för användning till ett kommande mobilspel för företaget Pyramid Beam Studios. Då spelet utvecklas för en mobil plattform så är det direkt relaterat till ljudkvalitet och storleken på dessa ljud. Förhoppningen är att detta arbete skall underlätta för ljuddesigners som vill utveckla ljud till mindre spel som kräver en viss typ av formatering. Förståelsen om ljud och hur kvalitet kan bibehållas trots en stor minskning i storlek samt statistisk kvalitet är en vital del till att kunna skapa kvalitativa ljudeffekter och musik till spel av denna typ. Slutligen så är det menat att kunna användas i ett skarpt projekt och tillföra nytta till framförallt spelbranschen i stort men även för andra som är intresserade av ljud överlag. Det här arbetet kommer gå in närmare på just samplingsfrekvens samt bit antal och vilken nivå av dessa visar sig vara kritisk för den uppfattade ljudkvalitén.

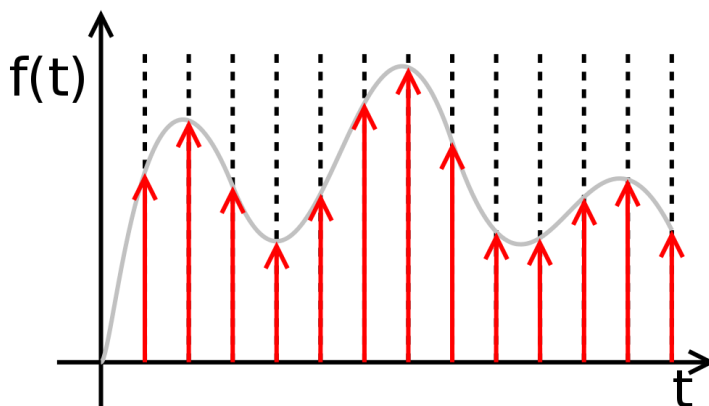
Historiskt så har vi haft persondatorer i ca 40 år vid tidpunkten av denna text och trots att det har funnits variationer i kraft och minne så har de i regel har ökat sedan dess. I slutet av 2000-talet så började dock det nya fenomenet "Smartphone" dyka upp. År 2007 så introducerades iPhone av Apple Inc vilket var startskottet för en revolution för den mobila marknaden och de spel som skulle följa därefter. Det som mobilerna förde med sig i denna nya era var även de restriktioner som man såg under 90-talet vad det gällde hårdvara och vad den klarar av att prestera i spelsammanhang. I och med att restriktionerna var så pass hårda, speciellt vad det gäller utrymme, så var man tvungen att anpassa sitt material som skulle användas i spel. För ljud så innebar det en rad olika problem, ett av de vanligare är att ljud ofta prioriteras bort om man jämför med t.ex. grafik, inte bara i gameplay men även vad det gäller utrymme (Belcher, 2010). Det man bör ha i åtanke är att ett spel till iPhone har en maxstorlek på 20 megabyte för att tillåtas nerladdning via 3G nätverk. Om detta mål inte uppnås så måste kunden ladda hem spelet via WiFi-nätverk vilket oftast betyder en minskning i försäljning för upphovsmannen (Belcher, 2010). Med dessa restriktioner i åtanke så innebär det att ljud måste vara så små storleksmässigt som bara möjligt medan de skall låta så bra som plattformen tillåter. Eftersom plattformen är mobil så innebär det ytterligare restriktioner än med en stationär plattform som används i hemmet. Högtalaren som finns på mobila enheter lämnar ofta mycket att önska vad det gäller frekvensåtergivning och ljudkvalitet, då dess minimala storlek inte kan ge önskvärt ljud så måste man jobba inom de ramar som högtalaren sätter upp. Det innebär bland annat att man bör skära bort frekvenser, särskilt de allra lägsta frekvenserna, som högtalaren inte kan spela upp då de kan förhindra de högre frekvenserna att komma fram (Belcher, 2010). I vissa fall så kan man även ha två olika uppsättning med ljud, en för högtalare och en för hörlurar. Detta är en möjlig lösning då hörlurar generellt har bättre frekvensåtergivning och kan ge lyssnaren en betydligt bättre upplevelse, speciellt om personen befinner sig i en miljö fylld av oljud. Det man bör ha i åtanke är att ljuden i sådant fall minst kommer ta lika mycket plats som de som är anpassade för högtalare, och därmed ytterligare minska utrymmet för olika varianter av ljud.

2.1 Termer

Det finns ett flertal olika termer som läsaren bör vara medveten om för att kunna förstå och ta till sig undersökningen samt resultatet. Vissa termer kommer utgöra en stor del av förklaringen över vilka ljud som är av hög eller låg kvalitet i undersökningen av detta arbete. Därför är detta en viktig del för läsaren som inte är medveten om hur digitala ljud fungerar. Förhoppningen är att dessa enkla och grundläggande förklaringar ska hjälpa till med förståelsen för undersökningsmetoden samt hur resultatet kan tolkas. Den här delen kommer att gå närmre in på sampling, kvantisering, bitupplösning samt en del övriga termer som kan vara av intresse.

2.1.1 Samplingsfrekvens

Samplingsfrekvens är ett mått på hur många gånger ett ljudkort samplar en analog signal i sekunden. Med sampling så menas det ett speciellt värde, som en ljudfrekvens, som läggs ut på en tidsaxel (Rubin & Tully, 1993). I detta fall så handlar det om en ljudsignal som samplas många gånger per sekund för att få en serie av värden som sedan kan tolkas av en dator för att återskapa signalen digitalt. Detta mäts i Hertz (Hz), till exempel 10 Hz läser av en analog signal 10 gånger per sekund. Ett av de vanligare värden man ser i samplingsfrekvens är 44.1 kHz, detta innebär då att signalen läses av 44 100 gånger per sekund, detta ger en digital signal som är mycket lik originalet och låter nästintill identiskt för det mänskliga örat. Just 44.1k Hz används på Cd-rom skivor som ett exempel på vilken ljudkvalitet det kan hålla. Vilket resultat sampling får beror givetvis på mer än bara samplingen i sig, även vilket bit djup filen håller.

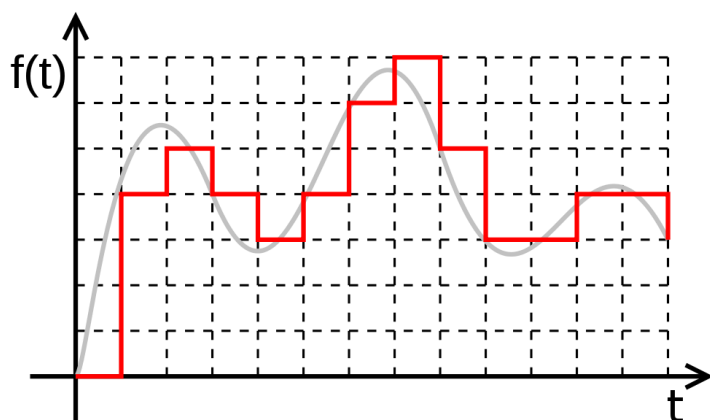


Figur 1 Här syns en samplad analog signal innan den har kvantiserats. Pilarna representerar de punkter där signalen samplas. (Wikipedia)

2.1.2 Bitdjup & kvantisering

Bitdjup i en digital ljudfil innebär att en algoritm avrundar frekvensen och amplituden till en av flera nivåer baserat på vilket bitdjup filen har (Rice, 2005). Detta kan jämföras med att man avrundar decimaler till heltal på en termometer. Detta innebär att ju högre bitdjup du har i en fil gör att du avrundar frekvens och amplitud till en som liknar originalet i det analoga ljudet. Detta kan liknas vid att du inkluderar halvtal i dina mätningar på en termometer och på så sätt får en mer exakt mätning.

Kvantisering går ut på att du tar en analog signal och ger den nivåer via bit djupet för amplituden på en frekvens som är hanterbara för en dator. Kort sagt göra om signalen till ettor och nollor (Rubin & Tully, 1993).



Figur 2: Slutligen så ser vi en kvantiserad digital signal där man tydligt ser punkterna där den samplats samt blivit kvantiserad. (Wikipedia)

2.1.3 Format

Det finns många olika format för ljudfiler och deras popularitet varierar kraftigt beroende på ändamål. Man kan dela upp formaten i två olika grupper, icke-förstörande komprimering och förstörande komprimering. Namnen är väldigt beskrivande då deras funktion i båda fallen reflekteras i det. Icke-förstörande komprimering är ett förlustfritt format som inte kompromissar ljudkvalitén när det sparas ner och bibehåller mycket kvalitet. Nackdelen med detta är att det även tar betydligt mycket mer utrymme än till exempel en MP3-fil av samma ljud. (Rice, 2005) påpekar skillnaden på datamängden som lagras i respektive format:

Each second of an uncompressed CD-quality recording requires 172K of storage. A lossy compression scheme offers reduced storage but at the cost of reducing the perceived audio quality. Popular lossy schemes, such as MP3, require only 4K to 40K of storage per second. However, the audio quality is acceptable to most listeners because the algorithms focus on discarding audio data that is not perceived. (Rice, 2005, A survey course on computer audio, s.120)

Med CD-kvalitet så menas det att det är ett icke-förstörande format som vanligtvis är Wav för en CD. Wav är ett av de vanligaste förlustfria formaten i världen medan MP3 är det vanligaste förstörande formatet, så att jämföra dessa två ger en bra bild över hur stor skillnad det blir på utrymmes användningen.

Som nämnt ovan är ett av de vanligaste förstörande formaten som finns MP3, vilket de flesta med största sannolikhet har använt vid något tillfälle. Det fungerar på ett sådant sätt att det väljer bort data ur en fil genom att sortera bort ljud som väldigt sällan skulle uppfattas av en människa för att skapa en trovärdig representation av originalfilen. Detta åstadkommes med hjälp av en algoritm som tar bort ljud som maskeras av existerande frekvenser och inte tillför speciellt mycket till ljudet i slutprodukten (Rice, 2005).

2.2 Tidigare forskning

För den här undersökningen användes litteratur från två olika områden, främst litteratur som fokuserar på den tekniska aspekten av ljudkvalitet. Även litteratur som behandlar ämnet psykoakustik, läran om hur människor uppfattar ljud, har kommit till användning. Valet av dessa två inriktningar gjordes då de behandlar de två aspekter i detta arbete som var helt vitala för att komma fram till ett resultat, samt att de har fungerat som stöd för arbetet i det avseendet att de innehåller fakta om de områden som undersöks. Dessa två områden, den tekniska aspekten av ljudkvalitet och psykoakustik kombinerat, som har legat som grund för arbetet och undersökningen.

Det finns mycket forskning inom området ljud men det som har varit mest relevant för denna undersökning har varit den tekniska aspekten av ljud och hur det fungerar. Anledningen till att arbetet har fokuserats på den tekniska delen av ljud och hur de fungerar har berott på den koppling arbetet har till företaget Pyramid Beam Studios. Arbetet är menat att användas som en typ av mall för ljudläggning för spel där en balans mellan ljudkvalitet och filstorlek är av yttersta vikt. Den tekniska ljudkvalitén skiljer sig från den upplevda i det avseendet att det är betydligt enklare att mäta eftersom det baseras på rena siffror snarare än hur en människa uppfattar ljudvågor. Fördelen med den tekniska fokusen tillsammans med den valda undersökningsmetoden för arbetet är hur tydliga resultat som genereras, samt hur det kan kopplas till och appliceras på ett spelprojekt.

Den litteratur som användes delades upp i två kategorier, en för de vetenskapliga texter som bildat grunden för arbetet. Den andra är övrig litteratur som använts vid projektarbetet eller metodvalet och undersökningsgenomförandet, samt som en faktabas för den grundläggande kunskap som förklarar hur ljud fungerar på ett grundläggande sätt. Anledningen till uppdelningen är för att klargöra vad som tillhör det som utgjort grunden för arbetet och vad som är granskade vetenskapliga texter. Litteraturen från båda kategorierna varierar i format och är hämtade från både böcker och publicerade artiklar på internet.

De vetenskapliga texter som använts innefattar främst den tekniska delen samt den psykoakustiska. Inkluderat i detta är även IEZA-modellen som beskriver ett sätt att kategorisera ljud. Den övriga litteraturen består av böcker samt artiklar som använts för att ytterligare beskriva den tekniska bakgrunden för arbetet samt fungerat som hjälpmedel för produktionen. Även metodboken som användes för val av undersökningsmetod innefattas av denna kategori.

2.2.1 Vetenskaplig litteratur

I *Spatial hearing: the psychophysics of human sound localization* så går Blauert (1996) in på hur människan faktiskt uppfattar ljud och vilka anledningar det finns till att vi hör saker på olika sätt. Blauert(1996)beskriver hur uppfattningsförmågan gradvis blir sämre med åren och vilket frekvensomfång som människan har i olika stadier i livet. Boken innehåller även delar som behandlar sättet människor hör, både det medfödda och det inlärd. *Spatial hearing: the psychophysics of human sound localization* är relevant för detta arbete i det avseendet att den tar upp hur människor hör, och hur vi lär oss att uppfatta olika typer av ljud och vad som filtreras bort.

Rice(2005) ger en kort introduktion till ljud i *A survey course on computer audio*, därefter ger han läsaren mer ingående kunskap om hur ljud fungerar med ett algoritmiskt synsätt där programmering av ljud och områdets bakomliggande kunskap behövs för att kunna jobba med det på en programmeringsnivå. Rices(2005) artikel skrevs i samband med att han undervisade i ljudprogrammering och han finner att hans elever blev mer bekväma och motiverade av undervisningen eftersom flera av dem visade ett intresse att fortsätta studera ljud. Artikeln är relevant i mitt arbete då den tekniska bakgrunden han beskriver används som underlag i undersökningen.

Tol, R., Huiberts, S. (2008) *IEZA: A framework for game audio* Användes för urvalet av ljud i undersökningen. Texten beskriver en typ av mall för att identifiera och kategorisera ljud för spel. Texten var relevant i det avseendet att det skapade en grund för valet av de ljud som användes i undersökningen och vilken funktion de skulle uppfylla i ett spelsammanhang. Det ansågs vara en viktig del av arbetet då det är baserat på en riktig produktion där de typerna av ljud som valdes för undersökningen kommer att användas.

Gabrielsson & Sjögren (1978) skriver i *Perceived sound quality of sound-reproducing systems* om upplevd kvalitet på ljud. Deras arbete skiljer sig från det här då det behandlar ämnet med en fokus på upplevd kvalitet på olika ljudsystem. De använder sig av liknande urvalsgrupp i deras undersökning som det här arbetet är tänkt att göra, d.v.s. en grupp av människor med en hög grad av medvetenhet av ljudkvalitet och en grupp som har mindre erfarenhet i ämnet. Även här finns dock skillnader, det här arbetet använder sig delvis av försökspersoner som inte har någon erfarenhet av ljuddesign och/eller Hi-fi medan (Gabrielsson & Sjögren, 1978) använder sig av två grupper med viss erfarenhet av ljud dock i olika grad. Deras resultat var intressant i relation till det här arbetet eftersom de fann att olika typer av lyssnare, d.v.s erfarna lyssnare jämför med noviser, har olika känslighet för ljud och upplevd kvalitet av ljud överlag (Gabrielsson & Sjögren, 1978).

2.2.2 Övrig litteratur

The Audible PC, (Rubin & Tully, 1996) är en bok som beskriver hur ljud och framförallt digitala ljud i datorer fungerar. De ger läsaren en grundkurs i hur ljudvågor kan konverteras till ettor och nollor för att sedan funka i en digital miljö. Rubin och Tully(1996) går även in på vad som kan gå fel när ljud konverteras, exempelvis aliasing, samt hur detta kan undvikas. Vad som är av störst intresse för detta arbete i *The Audible PC*, (Rubin & Tully, 1996) är förklaringarna om hur samplingsfrekvens fungerar och hur det tillsammans med bitdjupet är en vital del i hur digitala ljud fungerar, vilket är grundläggande kunskap för att kunna förstå forskningsresultat inom detta område.

Metodbok för medievetskap (Östbye m.fl. 2004) används i detta arbete för kunskap om val av metod i undersökningen, samt som en grund för varför just den metoden som valts är lämpligast. Östbye m.fl. (2004) beskriver ett flertal undersökningsmetoder och vad som kännetecknar dessa samt de olika metodernas fördelar och nackdelar. Boken användes främst till att komma fram till vilken typ av undersökningsmetod som var lämpligast för undersökningen. Utöver det så användes den för att säkerställa undersökningens kvalitet och tillförlitlighet genom att undvika eventuella misstag som beskrivs i boken.

Belcher (2010) skriver mer ingående om problemen som uppstår vid ljudläggning för en mobil plattform i *iOS Audio Design: What Everyone Needs To Know*. Han beskriver de restriktioner som finns när en ljudläggare arbetar mot en plattform som har små högtalare med begränsad frekvensåtergivning. Just i detta fall så använder Belcher (2010) av en iPhoneenhet som exempel, där han poängterar de begränsningar som existerar och hur han personligen arbetar för att ta sig runt dem eller jobbar med dem så de blir så bra som bara möjligt. Belchers (2010) text har varit viktig i arbetet då den har givit en djupare förståelse för de krav som ställs på en ljudläggare som arbetar mot en mobil plattform i stil med dagens smartphones. *iOS Audio Design: What Everyone Needs To Know* (Belcher, 2010) har även fungerat som en inspirationskälla för arbetet och hur det skall kunna tillämpas i ett skarpt projekt. Där är kunskap om kvalitet är av yttersta vikt för att få fram så bra ljud som det går utan att överskrida de begränsningar som plattformen sätter upp för ljuddesignern.

3 Problemformulering

I arbetet är problemformuleringen:

När upplever en människa en märkbar kvalitetsförlust i en digital ljudfil och vilken del samt nivå av det tekniska är kritisk för upplevelsen?

Med begreppet tekniska syftas det på saker som sampling rate, bit djup och format. Med begreppet psykoakustik menas det hur en människa upplever ett ljud, oavsett om det är en digital källa eller ej, men man bör ha i åtanke att i detta fall så kommer det uteslutande vara digitala ljudkällor som kommer att användas.

I skapandet av ljud mot en mobil plattform så uppstår det en rad problem som en ljuddesigner måste ha i åtanke för varje ljud som han eller hon skapar och förbereder för implementation. En av ljuddesignerns största fiende för att få fram ett kvalitativt ljud är det minimala utrymme som finns tillgänglig att arbeta med (Belcher, 2010). Detta är speciellt tydligt när man tittar närmre på utrymmet som tilldelas ljud i en applikation. Som det tidigare nämnts så är maxstorleken för en applikation som laddas hem via 3G 20 megabyte. Om detta överskrids måste kunden hitta ett WiFi-nätverk eller alternativt koppla in mobilen till datorn för att ha tillfälle att inhandla applikationen. Belcher (2010) menar då att nå 20 megabyte-gränsen är av yttersta vikt. Begränsningarna slutar dock inte där för en ljuddesigner, om gränsen nu är 20 megabyte så kan ljudet inte ta mer plats än ca 50% av det sammanlagda utrymmet. Det förutsätter att du får så pass mycket utrymme, grafik tenderar att ligga över ljud i prioriteringslistan, så slutligen kan man komma fram till att det finns maximalt 10 megabyte att jobba med i en typisk applikation.

Problemen med utrymmet slutar inte där, i exempelvis iPhonens fall så kan den inte spela upp mer än ett komprimerat ljud åt gången, vilket innebär att en stor majoritet av ljuden i en applikation måste vara okomprimerat i ett format, exempelvis Wav. Ett okomprimerat ljud tenderar till att ta upp stora mängder utrymme om man jämför med en komprimerad version som kan bli en runt femtedel av originalfilens storlek utan någon större märkbar skillnad vad det gäller kvalitet (Rice, 2005). För att kringgå detta problem är det en bra idé att välja ut ljud strategiskt och låta dessa vara komprimerade, som till exempel ett längre stycke musik (Belcher, 2010).

Förutom utrymme så bör man ha fler saker i åtanke när man skapar ljud till en mobil plattform. Speciellt den begränsade frekvensåtergivningen som de små högtalarna kan åstadkomma blir ett problem om det inte finns med i åtanke vid designen av ljudet. Därför bör en ljuddesigner vara väl medveten om vilka begränsningar som enheten/enheterna har som ljudet skall spelas upp på.

Alla de här problemen som har tagits upp kulminerar i vad det här arbetet handlar om, upplevd ljudkvalitet. Vissa av de beskrivna problemen kommer inte tas upp i den här undersökningen men läsaren bör vara medveten om problemen för att ge en bättre helhetsbild av resultatet och hur det förhoppningsvis kan användas i en produktion. Problemet som skall undersökas är hur en människa upplever en minskning av samplingsfrekvens eller bit djup och om det överhuvudtaget uppfattas av en människa eller ej. I takt med att mobiltelefoner utvecklas som medium så kommer ljuddesigners inte behöva arbeta mot de begränsningar som nu finns, men vid tidpunkten som den här undersökningen

utförs så är det ett faktum att det är nödvändigt att jobba mot de restriktioner som finns i samplingsfrekvens, bit djup och format för att få fram ett acceptabelt resultat.

3.1 Metodbeskrivning

Kortfattat så har undersökningen ett fokus på att testa ljud med olika tekniska parametrar inställda på olika former av kvalitet. Sedan genomförs ett blindtest med ett referensljud som försökspersoner av varierad teknologiskt kunnande får gradera.

Undersökningen innefattar den tekniska aspekten då olika former av kvalitet används. Detta kan yttra sig på olika sätt, så som samplingsfrekvens och bitdjup. Format kommer inte behandlas i denna undersökning då det finns för många olika format på marknaden för att kunna göra en genomgående och rättvis undersökning inom de tidsramar som finns. Det viktigaste och det som är essensen i undersökningen är i vilket skede människor uppfattar när antingen samplingsfrekvensen och/eller bitdjupet har försämrat ljudkvaliteten på en mängd varierade ljud.

En viktig del i undersökningen är miljön som försökspersonen kommer befinna sig i. För att försöka få en så rättvis bedömning som möjligt så ska miljön vara snarlik för samtliga försökspersoner och vara fri från störningsmoment (Östbye m.fl. 2004). Detta är för att få bort de variabler som kan förvränga resultatet. Förhoppningen är att om alla försökspersoner ges samma möjligheter och omgivning att uppfatta kvaliteten hos de ljud som används i undersökningen samt att den blir mer tillförlitlig.

3.2 Varianter av ljud

De ljud som används är varierade på mer än ett sätt. De varierar givetvis i kvalitet då det är det som hela undersökningen går ut på, men även i innehåll. Exempel på olika typer av ljud som används kommer synas som en lista senare i texten. I undersökningen är det viktigt att ha ljud med varierat innehåll för att se om vissa ljud är speciellt känsliga för kvalitetsförlust eller om det är gemensamt för samtliga typer av ljud. Ljud som innehåller information som till exempel mänskligt tal påverkas möjligen mer av kvalitetsförlust än en serie av pistolskott eller dyligt. Det här är en av de anledningarna till att undersökningen innehåller varierade försöksljud med olika typer av källor, både organiska i ursprung och ljud skapade med syntar.

En aspekt i undersökningen är att den är relaterad till arbetet kring mobilspel och hur ljud fungerar på en sådan typ av plattform. Därför är de ljud som testas av en sådan karaktär att de kan passa in i den typen av spel. De kommer dock i vissa fall vara längre och samtliga kommer inte vara anpassade efter någon speciell plattform. Anledningen till valet av ljud är att se huruvida de kommer påverkas likvärdigt av kvalitetsförlusten för att sedan hitta någon typ av optimalt ljud för mobila plattformar.

Läsaren bör ha i åtanke de olika nivåer av kvalitet som används. Skalan för samplingsfrekvens ligger mellan 44.1 kHz till 11 kHz, med 22 kHz och 16 kHz emellan. Bitdjupet på filerna är också varierade, från 24 bit till 8 bit med 16 bit mellan dem. Anledningen till att det inte används något högre, som till exempel 96 kHz och 32 bit, är för att dessa nivåer av kvalitet inte är relevant för en kvalitetsundersökning som inriktar sig på mobila plattformar då gränsen för dem generellt sett är lägre (Belcher, 2010).

Några exempel på vilka typer av ljud som används i undersökningen:

- Mänskligt tal
- Inspelat effektljud
- Inspelad explosion
- Syntbaserat effektljud
- Syntbaserad explosion

Mänskligt tal innehåller ett ljud, en kort fras på en mening som upprepas två gånger. Denna typ av mänskligt ljud är vanliga i spel och bör därför representeras i undersökningen för att se om det påverkas på samma sätt som ett effektljud.

Med inspelade ljudeffekter innebär det ett ”verkligt” ljud i den meningen att det är ett icke manipulerat inspelat ljud som du skulle kunna höra i utanför en digitaliserad värld. Detta kommer vara ett ljud av en låda som trillar ner på ett golv.

Den inspelade explosionen inkluderas då detta är ett mycket vanligt inslag i spel och bör finnas med i någon form.

De syntbaserade effektljud som kommer användas kommer vara en typ av ljud som producerats i en mjukvarusynt. De kommer vara ett laser ljud skapat med en mjukvarusynt. Det skall även kunna fungera som ett interfaceljud.

En syntbaserad explosion kommer användas då det är av intresse att se om det är någon direkt skillnad mellan en inspelad och en syntgenererad explosion i ett kvalitetsperspektiv.

3.3 Försökspersoner

De som ingår i undersökningen är en begränsad grupp av människor som passar in på olika parametrar. Detta för att undvika många olika och avvikande variabler som kan uppstå med en alltför bred urvalsgrupp. Förutom de två grupperna som används så har samtliga vissa gemensamma egenskaper; de är mellan 18 och 30 år gamla samt vara utan någon form av hörselskada eller nedsatt hörsel. Ålder påverkar hörsel i det avseendet att ju äldre en människa blir ju sämre blir den, framförallt så försämras förmågan att uppfatta höga frekvenser (Blauert, 1996). Relevansen i detta är framförallt exkluderingen av en eller fler äldre åldersgrupper då deras begränsade uppfattningsförmåga av höga frekvenser kan påverka slutresultatet. Man bör dock ha i åtanke att hörseln börjar försämras redan vid åtta års ålder (Blauert, 1996) så att ha en försöksgrupp med perfekt hörsel, om något sådant ens existerar, är inte bara orealistiskt utan även högst opraktiskt.

Försökspersonerna i undersökningen har en av två möjliga bakgrunder. En helt neutral sådan som inte har några direkta preferenser vad det gäller ljud i stort, exempelvis format, ljudutrustning, samplingsfrekvens och så vidare. Det medför även att det är mer kliniskt i själva utförandet och i miljön. En viktig del av undersökningen är att ta med människor som inte är tränade att höra efter ljudkvalitet och fel i ljudfiler. Detta för att få en mer realistisk grund i arbetet i det avseendet att de reflekterar en mer trolig kundbas för till exempel en mobil applikation.

Den andra gruppen som ingår i undersökningen är människor med någon typ av intresse eller bakgrund i ljud och/eller ljuddesign. Inkluderingen av denna grupp är ett sätt att eliminera eventuella förvrängningar av resultatet med en separat grupp. Det är även intressant att se om det är någon skillnad mellan grupperna i hur de uppfattar ljudkvalitén. En av anledningarna till att denna grupp inkluderades är att på Högskolan i Skövde finns det en utbildning med fokus på ljud vilket gör tillgängligheten av försökspersoner med just denna typ av bakgrund relativt enkla att få kontakt med. Det är också anledningen till att ha en separat grupp med försökspersoner utan någon ljudbakgrund då det lätt skulle kunnat ha smugit sig in en person med betydligt mer kunskap och tränad hörsel i undersökningen och gjort resultaten onormalt växlande.

3.4 Utrustning

Utrustningen som användes i undersökningen består av ett flertal olika komponenter. En dator för uppspelning samt hantering av ljud under experimentets gång. Ett externt ljudkort för att säkerställa ljudkvaliteten som försökspersonerna utvärderar. Samt studio monitor-hörlurar för att få ett så klart och platt ljud som möjligt. Alla komponenter tillsammans är menade att säkerställa ljudkvalitén som helhet.

En kort lista på samtliga komponenter som används vid undersökningen och som inte ändrats under undersökningens gång.

- Lenovo IdeaPad G560 Laptop
- MidiTech AudioLink II USB Ljudkort
- Superlux HD 668B Studio monitor hörlurar

Samtliga försökspersoner använde sig av denna utrustning för att undvika varierande resultat beroende på annat än förmågan hos försökspersonerna. Miljön där undersökningen tog plats var av en snarlik karaktär, det var en tyst plats med så få störande ljud och moment som möjligt för att ge försökspersonen en så god chans som möjligt att uppfatta eventuell kvalitetsförlust i ljudfilerna som användes.

3.5 Undersökningen

Undersökningen gick till på ett strukturerat sätt med hjälp av enkäter för att undersöka hur försökspersonerna uppfattar de ljudfiler som spelas upp. Metoden som används för undersökningen är en kvantitativ metod som Östbye m.fl. (2004) beskriver ingående i *Metodbok för medievetenskap*. En fördel med metoden är att den data som samlas in blir rena nummer och är därmed lättare att sammanställa än till exempel intervjudata av en kvalitativ forskning. En ytterligare fördel med en kvantitativ metod är mängden av data den genererar, det blir måttliga mängder data som inte tar för lång tid att sammanställa, i jämförelse med den kvalitativa metoden som kan vara svår att få fram jämförbara resultat i. Utöver de fördelar som nämnts så är den här metoden också mycket tydlig i sin presentation i utvärderingssyfte. Avsikten med detta i undersökningen är att ge läsaren en tydlig representation av resultatet så de kan kritiskt granska all data.

De nackdelar som finns med den här typen av undersökning är att den inte är fullt realistisk då undersökningen skedde i ovanligt bra förhållanden för lyssnande. Man bör även ha i åtanke att försökspersoner med hörfel, exempelvis tinnitus, är helt uteslutna från undersökningen vilket gör den än mindre realistisk och kan inte ses som generell för befolkningen, men den blir betydligt mer konsekvent. Problemet med att undersökningen blir klinisk har tagits i åtanke och undersökaren är väl medveten om detta. Dock så anses denna metod att vara den mest lämpliga för denna typ av undersökning då den data som uppstår blir lätthanterlig och ofta mycket tydlig. En aspekt som togs med i valet av undersökningsmetod var även de tidsramar som finns för arbetet. En kvantitativ undersökning innebär en kortare utvärderingsperiod vilket är en fördel då det finns strikta tidsramar som måste följas.

3.6 Utvärdering

Utvärderingen av resultatet skedde med data som sammanstälts från enkäter som försökspersonerna fått svara på medan de lyssnat på olika ljud för att sedan gradera dessa på en femgradig skala, en så kallad Likertskala. Fördelen med denna skala är att 5 punktskalan anses vara tillräcklig för att få fram ett sanningsenligt resultat samt att den mängd med data som skapas i samband med undersökningen blir mer lätthanterlig (Bryman, 2001). Skalan ska bestå av olika nivåer av uppfattad kvalitetsförlust. Resultatet av enkäterna skall sedan sammanställas till stapeldiagram beroende på ljud samt teknisk kvalitet. I slutet skall de även jämföras med varandra för att se huruvida det är skillnader på olika typer av ljud och vilket grupp som utmärkte sig eller om de är nästintill identiska.

Enligt Gabrielsson och Sjögren (1978) så visade det sig att försökspersoner med en bakgrund i Hi-fi och som hade ett aktivt intresse för kvalitativt lyssnande märkte lättare av oljud än deras motpart i undersökningen. Detta är något som diskuteras i resultatet av undersökningen då användandet av två olika grupper kan ge en bra mittpunkt att lägga kvalitén på i en produktion för att tillfredsställa båda målgruppernas förväntningar på ljudkvalitet. Förutom den praktiska aspekten så är det även intressant för själva utvärderingen för att se om den når ett liknande resultat som Gabrielsson och Sjögrens (1978) undersökning gjorde.

Frågeställningen besvarades med den data som producerats av undersökningen och sammanstälts till tydliga diagram som sedan jämförts med varandra. Analysen av resultatet innefattar en jämförelse av de två grupper av försökspersoner som är med i undersökningen. Detta för att finna en möjlig skillnad mellan de två grupperna och hur pass stor den skillnaden är. En analys på de olika typerna av ljud har genomförts, vilken används för att se om det finns någon skillnad när man har samma grad av teknisk kvalitet men olika typer av ljud. Ett exempel på detta är om det är någon skillnad på hur försökspersonerna uppfattar inspelade ljud eller syntbaserade ljud.

4 Genomförande

Materialet som användes för undersökningen är ljud av olika ursprung som skulle kunna användas i ett spel. Varje ljud som användes har ett kontrollljud på en identisk kvalitetsnivå, i detta fall så har samtliga ljud en samplingsfrekvens på 44.1 kHz och 24bitdjup. Valet av just denna nivå grundar sig i att den håller god ljudkvalitet samt att ljudkvalitet för mobilspel generellt sett är mycket lägre (Belcher, 2010). Därmed så är en högre kvalitet än så är inte nödvändig eller relevant för undersökningen. Ljuden som använts i undersökningen har skapats på varierande sätt. Runt hälften är av ett syntbaserat ursprung och resten är inspelade med en portabel inspelare tillsammans med en dynamisk mikrofon. Anledningen till användandet av båda typer av ljud är för att testa om syntbaserat material och inspelat material skiljer sig i kvalitetsaspekten. Arbetet relaterar även till företaget Pyramid Beam Studios i det avseendet att se vilken typ av ljud som lämpar sig främst till mobilspel. Då kvaliteten på ljud i en mobil plattform måste komprimeras hårt för att få plats på en mobil, så ligger det i företagets intresse att veta om vissa typer av ljud fungerar bättre hårt komprimerade än andra.

Valet av ljud som används i undersökningen baserades därför på typer av ljud som troligen skulle kunna hittas i ett flertal spel. Detta kopplades även till den så kallade IEZA-modellen (Tol, Huiberts, 2008) där författarna skiljer på ljud som existerar inom spelvärlden, diegetic sounds. Samt non-diegetic sounds, ljud som kommer utanför spelvärlden. Diegetic sounds är saker som hörs i spelvärlden, till exempel; explosioner, tal och kolliderande objekt i spelvärlden. Non-diegetic sounds är saker som UI ljud samt musik, ljud som ger ett spel en speciell stämning. Relevansen för det här arbetet ligger i att mobilspel tenderar till att ha en betydligt mer begränsad ljudbild än en högbudget PC eller konsol-titel, detta beror på ett flertal faktorer som budget och utrymmesbegränsningar och så vidare (Belcher, 2010). Detta innebär att ljud för mobilspel är betydligt mer sparsamt använda och fokus ligger på de viktigaste ljuden för att spelet skall fungera. På så vis influerade IEZA-modellen valet av ljud för undersökningen.

De ljud som ligger i fokus för spelet *So gravity!* som Pyramid Beam Studios utvecklar, och som är relaterat till arbetet är huvudsakligen skapade för att få en fungerande ljudbild utan att överskrida de begränsningar som finns till den mobila plattform som spelet är avsedd för. Det innebär att majoriteten av de ljud som spelet är i behov av är av en diegetic sound-karaktär vilket medför att en majoritet av ljuden som togs fram för undersökningen räknas in i en effektljudskategori. Ljud som tal, inspelat effektljud, inspelad- samt syntbaserad explosion är inräknade i dessa. Ytterligare ett ljud som har egenskaper för att passa in i både diegetic och non-diegetic ljud användes för undersökningen, vilket blev det syntbaserade effektljudet. Detta är avsett för att kunna användas både som ett UI ljud samt som ett vanligt effektljud. Valet av just dessa ljud grundar sig i att de skall ha ett riktigt användningsområde för spelproduktion snarare än att fylla samtliga olika kategorier i IEZA-modellen, men ändå att ha den i åtanke och inkludera ljud som varierar i innehåll och funktion.

Även psykoakustiken influerade valet av ljud för undersökningen. Människan uppfattar tal i en annan grad än ljud som inte innehåller någon information (Blauert, 1996). För att beskriva detta kan man använda den så kallade *Cocktail-party effekten*, vilken innebär att en människa är kapabel till att föra en konversation med någon annan och höra vad den säger även om båda talarna är omringade av ljud från en mängd olika konversationer som blir till störningsljud (Blauert, 1996). Ovanstående blir av intresse för undersökningen då försökspersonerna möjligtvis finner att tal är mer eller mindre mottagligt för kvalitetsförlust. Detta kan resultera i att en ljudläggning för ett spel kan behöva justera kvaliteten för just tal beroende på resultatet.

4.1 Praktiskt genomförande

De syntbaserade ljuden har skapats med hjälp av programmet Propellerhead Reason 5 (Propellerhead Software, 2010). Valet av detta program gjordes för dess starka fokus på mjukvarusyntar av hög kvalitet. Den syntbaserade explosionen och det syntbaserade effektljudet skapades med mjukvarusynten Thor. Tillvägagångssättet var okomplicerat då Thor har många så kallade patchar. Patchar är förprogrammerade ljud som kan laddas in för att snabbt få ljud med varierande funktion. Dessa fungerade som en utgångspunkt för arbetet och användes som en grund för ljuden som användes i undersökningen. De olika patcharna behövde minimalt med förändringar för att bli av en önskvärd karaktär för undersökningen och själva arbetet låg mer i att spara ner dessa filer i respektive kvalitetssteg som användes för undersökningen.



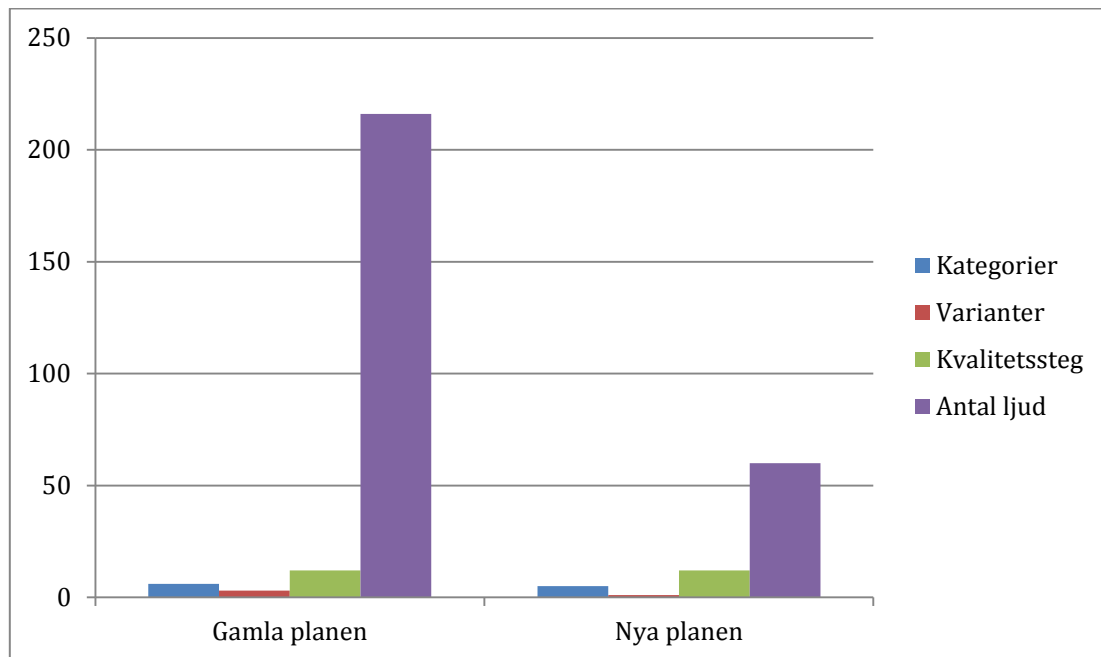
Figur 1 Mjukvarusynten Thor i Reason 5 (Propellerhead software, 2010)

Det inspelade materialet spelades in med hjälp av en Roland R-05 portabel inspelare och en Superlux D108A dynamisk mikrofon. Läsaren bör ha i åtanke att dessa två produkter inte är av den högsta kvalitet vilket möjligtvis bidrar till sämre kvalitet i grundmaterialet. Detta bör dock inte ha någon inverkan på slutresultatet då undersökningen berör kvalitetsförlust från grundmaterialet och inget annat. Materialet spelades in med en sampling frekvens på 48 kHz och 24 bitdjup och sparades sedan ner i de kvalitetsstegen som används för undersökningen. Innan det sparades ner så normaliserades samtliga ljud och behandlades med klippning för att få bort oönskat material. Det inspelade effektljudet gick även genom ytterligare en process då de enskilda ljuden från inspelningarna inte var bra nog så klipptes två av dessa ihop för att ge ett mer effektfullt och varierat ljud.

Ljudet med högst kvalitet, det vill säga kontrollljudet, exporterades ut med ett bitdjup på 24 bit och en samplingsfrekvens på 44.1 kHz. Därefter så importerades dessa ljud in i programmet Cubase 5 (Steinberg GmbH, 2009) där de sparades ner i de lägre kvalitetsstegen som används i undersökningen. Samma metod användes till samtliga ljud, inspelade som syntbaserade. Valet av att använda samma program för sparandet av samtliga ljudfiler grundade sig i att ju färre variationer det finns desto bättre för undersökningens tillförlitlighet. Variationer i komprimeringen, hur än små, bör undvikas i undersökningen då det kan skapa ytterligare variabler att ha i åtanke när resultatet skall utvärderas. Av just denna anledning gavs samtliga ljud samma behandling i programmet, samtliga ljud normaliserades på samma sätt med en övre gräns på -1 dB innan de sparades ut i de kvalitetsnivåer som används i undersökningen.

Tidigt i produktionen av ljuden så blev det tydligt att omfattningen av de ljud som skulle vara med i undersökningen var alltför stor för de tidsramar som arbetet var tvungen att hålla. Därför genomfördes det ett urval av ljuden som slutligen kom med i undersökningen. Till en början valdes sex kategorier med ett ljud per kategori snarare än tre ljud. Till exempel så skulle mänskligt tal kategorin ha tre olika fraser som skulle läsas upp. Detta ändrades till endast en fras då mängden data att sammanställa med tre olika fraser per kategori hade riskerat att deadline för undersökningen inte hade kunnat mötas. För att ge en överblick på mängdskillnaden så gjordes följande grafiska representation:

Tabell 1 Ljudmängd från tidigt till sent planeringskede



Skillnaden på antal ljud var 216 ljud enligt den gamla planeringen mot 60 ljud enligt den nya planeringen. Det som gjorde mängden mer hanterbar var minskandet av varianterna för de olika kategorierna. Senare i genomförandet av ljudproduktionen upptäcktes ytterligare ett problem. I detta fall så var problematiken det syntbaserade talet som visade sig vara mer av en utmaning än vad som tidigare antogs. Problemet grundar sig i att det finns mycket få program för röstsyntes som tillåter sparande av högkvalitativa filer till en dator. Det ledde till problemet med att få ett högkvalitativt grundljud att spara de resterande ljud från. I slutändan ledde detta till att det syntbaserade talet helt enkelt fick tas bort ur undersökningen så att fokus fick ligga på de ljud som faktiskt kan användas. Om detta dock sätts i perspektivet att ljuden skall kunna tänkas representera ljud från ett spel av någon form så hör röstsyntes inte till vanligheterna. Röstsyntes var mest intressant för undersökningen i det avseende att det eventuellt kan finnas en skillnad mellan vanligt tal och röstsyntes i ett kvalitetssammanhang.

4.2 Förstudie

Innan undersökningen startades så gjordes valet att genomföra en förstudie. Anledningen var till största del att snabbt få reda på tidsåtgången som undersökaren samt försökspersonen kan förvänta sig av en undersökning. Detta gjordes för att kunna planera för den riktiga undersökningen samt få en uppfattning om tidsåtgången som krävs för att kunna genomföra hela projektet. Ytterligare en fördel med förstudien var att hitta eventuella problem med metoden innan den riktiga undersökningen började (Östbye m.fl., 2004). Samt möjligheten att korrigerar dessa utan att överskrida tidsramarna för projektet.

Förstudien var relativt liten, även för en förstudie, med endast två försökspersoner. Målet med studien var som tidigare nämnt att få en uppfattning om tidsåtgången snarare än att hitta stora mängder fel i metoden som användes även om det också inkluderades i förstudien. Förstudien gick till på samma sätt som planerat för den riktiga undersökningen dock med ett annat val av lokal och till viss del utrustning. Utrustningen och lokalen ansågs inte vara

väsentligt för förstudien då fokus låg på att försöka hitta brister i metoden samt att få en tidsuppskattning. De två försökspersonerna som användes har ingen bakgrund i ljudläggning eller Hi-fi, det vill säga de har ingen erfarenhet av ljudläggning eller speciella preferenser vad det gäller ljudutrustning eller liknande. Utöver tidsåtgången så testades även deras uppfattningsförmåga av förlust av ljudkvalitet. Det togs även tillfälle att se huruvida undersökningsmetoden fungerar och om den krävde justeringar.

Utförandet av förstudien gjordes på samma sätt som planerades för den riktiga undersökningen med de undantag som beskrevs tidigare. Försökspersonerna i förstudien använde sig av samma enkäter som de som planerats ingå i den riktiga undersökningen och fick lyssna på exakt samma ljud. De fick hörlurar och de papper som innehåller de femgradiga skalor som ljuden ska bedömas med. Testet har fem papper med elva ljud som skall bedömas, tolv ljud spelas totalt varav ett är kontrolljudet. Tidsåtgången mättes med en Nokia Lumia 800 mobiltelefon med Windowsphone 7.5 operativsystem. Nedan syns tidsåtgången för båda försökspersonerna samt medelvärdet för tidsåtgången. Målet för tiden var runt 15 minuter per försöksperson vilket det tycks vara nära. Tidsåtgången förväntas att gå ner ytterligare när undersökningen blivit rutin.

Tabell 2 Tidsåtgång vid förstudie

Försökspersoner	Tidsåtgång
Försöksperson 1	17 minuter 28 sekunder
Försöksperson 2	14 minuter 26 sekunder
	Medelvärde: 16 minuter 17 sekunder

Till en början var det planerat att spela upp kontrolljudet endast en gång, detta visade sig göra det svårt för försökspersonerna att bedöma om ljudet hade förändrats. En försöksperson i förstudien informerade om att det gick för lång tid mellan då kontrolljudet hördes och de senare ljuden som spelades upp i undersökningen. Konsekvensen av det var att en märkbar förändring genomfördes, vilket var att spela upp kontrolljudet med jämna mellanrum så försökspersonen hade det färskt i minnet. Då kan en bättre bedömning göras om huruvida ljudet hade förändrats från originalet. Båda försökspersoner visade resultat som är att förvänta vad det gäller kvalitetsuppfattningen. När samplingsfrekvensen hålls på minst 22 kHz och bitdjupet på minst 16 så tycks de uppleva att kvalitén var oförändrad eller bara med en mycket liten förlust i kvalitet. När ljuden gick ner på 16 kHz i samplingsfrekvens varierade resultatet beroende på vilket ljud som spelades upp. I de flesta fall så ansågs det vara en klar förlust i kvalitet medan i andra så var det inte lika märkbart. När ljuden gick ner på 8 bit så reagerade båda försökspersonerna på en märkbar förlust i kvalitet vilket förväntades.

Förstudiens syfte var att få fram en tidsuppskattning samt en granskning av metoden och hur den fungerade för denna typ av undersökning. De exakta resultaten kommer inte att redovisas i arbetet då detta inte var målet med förstudien. Det har heller inte någon relevans för den senare undersökningens resultat.

5 Analys

Undersökningens genomförande skedde med hjälp av en kvantitativ metod, detta innebär att resultatet inte kommer att innehålla personliga reflektioner från försökspersonerna om kvalitetsförlust i ljudfiler. Metoden kommer istället användas för att analysera de resultat som undersökningen gett med hjälp av den data som samlats med hjälp av enkäten. Sammanställningen av data skall förhoppningsvis ge läsaren en god bild över när en människa upplever förlust av kvalitet samt vad det är som är mest vitalt för en positiv upplevelse. Viss statistik och tabeller kommer att användas för att presentera resultatet och kopplas till den frågeställning som skall besvaras i respektive underkapitel. Samtliga resultat skrevs in i ett bifogat Excel dokument där undersökningsresultaten kan läsas i sin helhet, det kommer även finnas tillgängligt i appendix A. Följande underkapitel kommer att behandla resultatet av undersökningen och besvara de frågeställningar som ställdes tidigare i arbetet.

Det som har analyserats är de mönster som märktes av i undersökningens gång, samt om de överförts till slutresultatet. Speciellt intressant är det att se huruvida mönstren var väntade, särskilt för de två undersökningsgrupper som använts i undersökningen. Till exempel om ljuden uppfattades på ett likartat sätt av båda grupperna eller om en grupp var mer benägen att ge ljuden en hårdare bedömning än den andra. Med hårdare bedömning menas att ljudet bedömdes att ha förlorat mer kvalitet än förväntat. Något annat som är av intresse är att se huruvida grupperna bedömde olika typer av ljud på olika sätt, om en viss typ av ljud var tydligare i sin förlust för en grupp men mindre tydligt för den andra.

Ytterligare frågor som skall besvaras med hjälp av undersökningsdata inkluderar även huruvida ljud från olika källor påverkas på ett annorlunda sätt i jämförelse med varandra. Ett exempel är om ett syntbaserat ljud låter bättre vid lågt bitdjup eller samplingsfrekvens än ett ljud som är inspelat från en naturlig källa med en mikrofon. Det blir av intresse för undersökningen då det är möjligt att vissa ljud är att föredra över andra vid en ljudläggning mot en mobil plattform som kräver en hård komprimering.

Slutligen så skall även huvudfrågan i undersökningen besvaras. Vilken del och nivå av bitdjup och samplingsfrekvens är viktigast i ett kvalitetssammanhang, samt vilken nivå av dessa två som krävs för att ge ljud en acceptabel kvalitet. Detta kopplas även till den tidigare punkten angående om det är någon skillnad mellan ljud av olika ursprung som syntbaserade ljud eller inspelade ljud. Ytterligare en fråga som besvaras är om ljud som innehåller information som mänskligt tal påverkas i en starkare grad av bitdjupsförminskning eller en lägre samplingsfrekvens. Denna del som tidigare nämnts är huvudfrågan vilket innebär att fokus kommer ligga på att besvara den på ett tydligt och övergripande sätt. Delen som behandlar just detta tar upp majoriteten av analysen men kopplas också till de övriga punkterna.

Undersökningen resulterar slutligen i en kvalitetsnivå där kvalitet och låg filstorlek kan kombineras och användas för ljudläggning till exempelvis mobilspel. Nivån skall kunna anpassas för de olika typerna av ljud som används i undersökningen för att se huruvida vissa ljud kan behöva högre kvalitet för att låta bra medan andra ljud kan sparas i en lägre kvalitet för att spara mer utrymme. Förhoppningsvis så resulterar det i en god överblick i vilken situation man bör använda olika typer av ljud och vad de lämpar sig för i ett kvalitetsperspektiv.

5.1 Mönster

Tidigt i undersökningen så märktes det ett flertal tendenser för de två grupper som medverkade i undersökningen. Båda tycks ha en liknande syn på kvalitet och när den försämras, speciellt när undersökningen gick ner på 8 bit ljud så reagerade båda försöksgrupper på ett enhetligt sätt med ett fåtal undantag. Det fanns även vissa skillnader mellan ljud av ett syntbaserat ursprung och de som var inspelade. I vissa fall så visade det sig att ljud med ett syntbaserat innehåll klarade av en större kvalitetsförlust än de som var inspelade. Hur och vad som kunnat orsaka detta kommer analyseras vidare i ett senare skede av detta kapitel. Tydliga tendenser i undersökningen observerades i ett flertal fall, vissa förvånande, andra helt förväntade. Allra tydligast är bedömningen av 8 bitdjup ljuden, samtliga ljud fick nästan alltid det sämsta betyget med ett fåtal avvikelser. Avvikelserna syntes tydligast i det syntbaserade effektljudet, vilket lät förvånansvärt klart vid lågt bitdjup. Först när samplingsfrekvensen blev låg så var det ett nästintill enhetligt lågt betyg. Det syntes även i vissa fall en tendens att gruppen utan en bakgrund i ljud och/eller ljudläggning bedömde de ljud som var med i undersökningen mindre hårt än de som var med i gruppen utan någon bakgrund i ljud eller ljudläggning. Det varierade dock beroende på vilken nivå av kvalitet som ljudet spelades upp för. Vid de lägre stegen av kvalitet för vissa av ljuden finns det fler positiva bedömningar för ljudkvaliteten i ljudgruppen än i gruppen med försökspersoner utan erfarenhet av ljud. Vanligast var dock att gruppen utan erfarenhet bedömde ljud mindre hårt.

Ytterligare ett mönster som syntes relativt snabbt var ett tydligt negativt intryck av ljud med låg samplingsfrekvens. Efter att ha gått under 22 kHz i samplingsfrekvens var bedömningen ofta mycket negativ. Även vid lägre bitdjup märktes det en viss acceptans av de ljud som sparats med en hög samplingsfrekvens. Dock vid 8 bitars bitdjup blev försökspersonerna kritiska av ljudkvaliteten oavsett samplingsfrekvens, vilket också var väntat. Till skillnad från bitdjupet så var försökspersonerna till stor del förutsägbara i deras bedömning av minskandet av samplingsfrekvensen. När samplingsfrekvensen var mellan 44.1 kHz och 22 kHz ansågs ljudkvaliteten vara relativt god, det stämmer överens med vad Belcher (2010) skriver i sin artikel. Viss förlust märktes men ingen av försökspersonerna ansåg att den var mycket märkbar i något av fallen förutom när ljuden hade sparats i 8 bitars ljud vilket bidrog till en mycket sämre kvalitet, mer än vad samplingsfrekvensen gjorde i de fallen. Även vid 16 kHz samplingsfrekvens så ansåg ett flertal försökspersoner att ljudet hade försämrats i en liten grad. Dock med många variationer i vilken grad det hade försämrats, samt även varierat beroende på vilket ljud de lyssnade på.

Vid försämring av bitdjup märktes det tydligt när försökspersonerna ansåg att kvaliteten hade försämrats. Vid en hög samplingsfrekvens märktes det mycket liten skillnad mellan 16 och 24 bitar enligt undersökningen. Först när bitdjupet sänktes till 8 blev det en markant skillnad på upplevelsen enligt försökspersonerna. Ur en kvalitetssynvinkel tycks bitdjupet påverka ljudupplevelsen i en liten grad så länge den hålls över 8 bitar vilket var något som påverkade ljudkvaliteten på ett mycket märkbart sätt. Hade det inte varit för undersökningens begränsade tidsramar hade ytterligare kvalitetssteg i bitdjup varit intressant för undersökningen. Speciellt ytterligare ett steg mellan 16 och 8 bitar som till exempel 12 bit för att se om det finns en bättre lägsta nivå än 8 bitar, vilket nästintill ingen i undersökningen ansåg hålla god kvalitet. 8 bitar gränsade till att vara nästintill ointressant för undersökningen då kvalitetsförlusten blev så stor att aliasing uppstod. Till och med i den grad att det skapade långa pipljud i vissa ljudfiler vilket gjorde kvalitetsförlusten både uppenbar samt oanvändbar

för en riktig produktion. Det i sig leder till att 8 bitars ljud skulle utan någon större förlust för en möjlig framtida undersökning uteslutas för att vidare studera ljudkvalitet.

5.2 Statistik

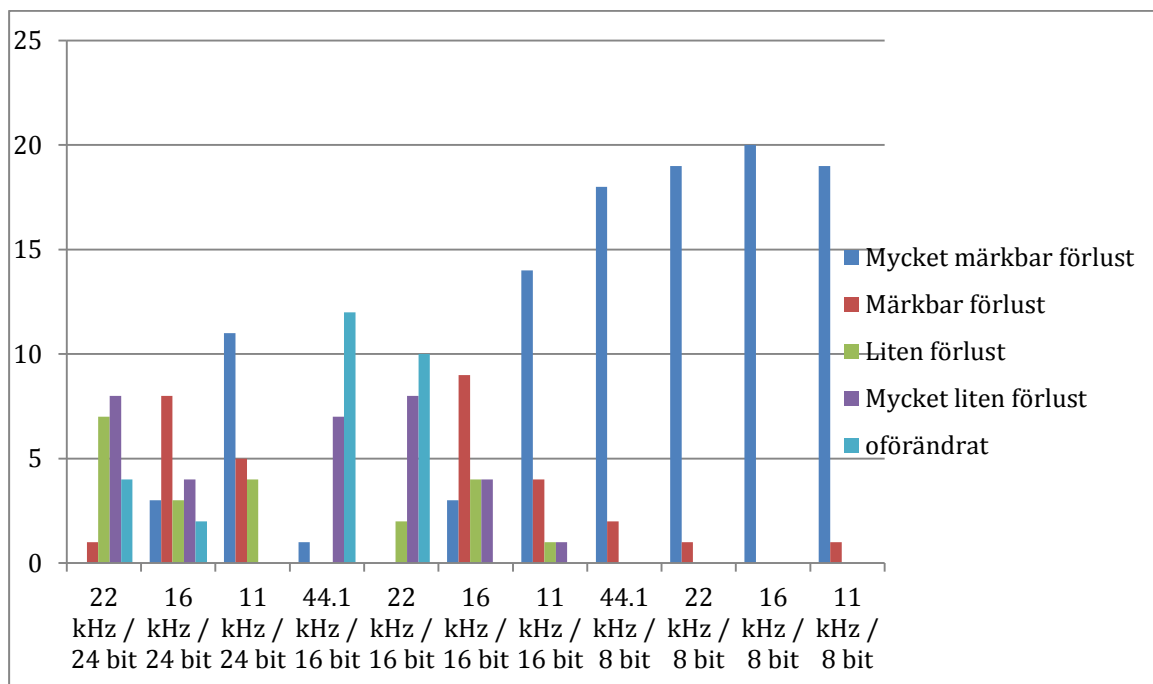
I den statistiska analysen av undersökningen finns det en stor mängd siffror att presentera för läsaren. För att undvika en överväldigande mängd tabeller och data så kommer inte all data presenteras i detta kapitel men det finns tillgängligt för läsning i sin helhet i appendix A. Data presenteras samt analyseras för att besvara de frågeställningar som detta arbete baserats på i olika segment. Därefter följer sammanfattningen av den analys som gjorts på den data som samlats in.

Segmenten behandlar de frågeställningar som undersöks och skall representeras av relevant data som förhoppningsvis skall ge en bra bild över hur människor uppfattar kvalitetsförlust i de ljud som finns med i undersökningen. Statistiken delas upp i olika grupper om tre, först den grupp som har en bakgrund i ljudläggning följt av gruppen utan någon bakgrund eller preferens i ljud. Till sist slås båda gruppernas resultat ihop för att ge en helhetsbild av vad försökspersonerna svarade på undersökningen. Utöver den statistiska analysen av försökspersonernas uppfattningsförmåga baserat på de enskilda ljuden i undersökningen så har det även jämförts med olika typer av ljud. Det vill säga om de uppfattade någon skillnad på ett syntbaserat ljud eller ett inspelat. Samma procedur används för att analysera detta, först presenteras de två olika gruppernas enskilda resultat per ljud. Sedan följer en sammanslagning av deras resultat för att se huruvida resultaten är enhetliga med vad de två grupperna tyckte enskilt. Det analyseras även om ljud som tal är mer eller mindre påverkat av kvalitetsförlust jämfört med de andra ljuden i undersökningen. Det följer givetvis samma form som de tidigare analyserna med enskilda grupper och till sist en sammanslagning av deras resultat.

5.2.1 Ljud: Inspelad explosion

Enligt den statistik som undersökningen genererat kan man tydligt se att grupp 1 som är den grupp som har en bakgrund i ljudläggning eller Hifi bedömde att ljudet hade försämrats i större grad än grupp 2 som är gruppen utan erfarenhet av ljud. Men ju längre ner ljudkvalitén sjönk så blev de mer och mer ense om hur förändrad den var, med vissa undantag. Vid 8 bitars djup var båda grupper nästan helt överens om att kvalitén hade fått en mycket märkbar försämring vilket inte var helt förvånande då det vid 8 bitdjup uppstod det mycket sprakande samt missljud som påminner om ett bitcrusch-filter. Skillnaden på ljudkvalitén blev markant mellan 16 bitar och 8 bitar vilket förklarar hur överens båda grupperna var med sin bedömning. Vid 16 bitar var grupperna dock mindre överens vilket kan förklaras med att ljudkvalitetsskillnaden från referensljudet på 24 bitdjup inte är speciellt stor om en hög samplingsfrekvens ackompanjerar 16 bitars ljuden. Vid 16 bitar så började åsikterna variera i större grad än de på 8 bitdjup, båda grupperna ansåg att ljudet var nästintill oförändrat vid 44.1 kHz med 16 bitdjup, dock så drog åsikterna isär när samplingsfrekvensen sjönk. Vid 22 kHz så bedömde grupp 1 ljudet hårdast medan vid 16 kHz så var det grupp 2 som bedömde ljudet hårdast. Det var dock tydligt att samtliga deltagare ansåg att ljudet hade förändrats på ett negativt sätt och att det var märkbart i en viss grad.

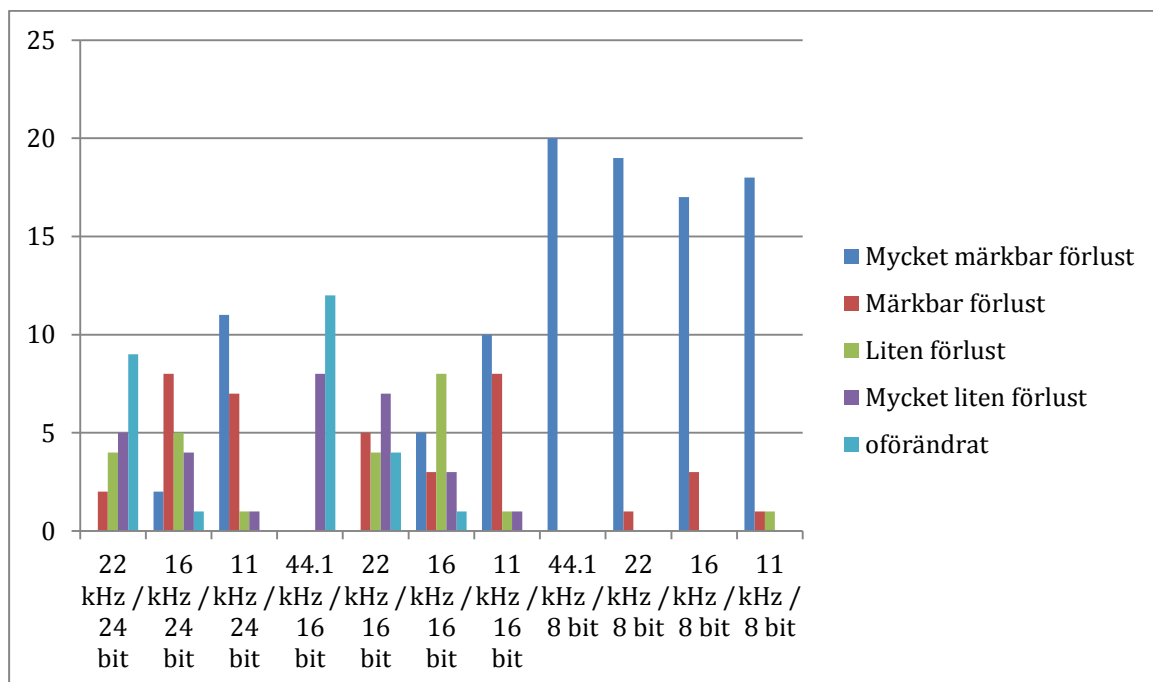
Tabell 3 Resultat för samtliga försökspersoner för Inspelad explosion.



5.2.2 Ljud: Syntbaserad explosion

Det syntbaserade ljudet bedömdes, precis som den inspelade motsvarigheten, på ett varierat sätt. De ljud som hade 24 bitdjup bedömdes hårdare av gruppen med erfarenhet av ljud än den utan erfarenhet. Dock bedömdes samma ljud med en viss spridning från båda grupperna vilket tyder på att det var svårbedömt för samtliga deltagare i undersökningen. Vid en nedgång till 16 bitdjup så var båda grupperna relativt jämlika i sin bedömning, som med det motsvarande inspelade ljudet. Vid en hög samplingsfrekvens, speciellt vid 44.1 kHz, ansåg båda grupperna att ljudet antingen var oförändrat eller att de upplevde en mycket liten förlust. Vid 22 kHz delade sig åsikterna avsevärt både inom grupperna men även mellan grupperna. Åter igen så bedömde grupp 1 ljudet hårdare än grupp 2, dock delade sig åsikterna åt och en viss spridning av bedömningen observerades. Grupp 2 bedömde ljudet på ett liknande sätt med spridda åsikter om huruvida ljudet hade förändrats lite eller inget alls. Båda grupperna tycktes ha svårt att bedöma detta ljud, och speciellt ljudet med 22 kHz samplingsfrekvens och 16bitdjup där åsikterna var mycket spridda. Det observerades att försökspersonerna bedömde just det kvalitetssteget mycket varierat i samtliga ljudtyper. Vid 8 bitdjup så observerades ett förväntat och liknande resultat i jämförelse med de andra ljudtyperna. En klar majoritet ansåg att ljudkvaliteten hade en mycket märkbar förlust.

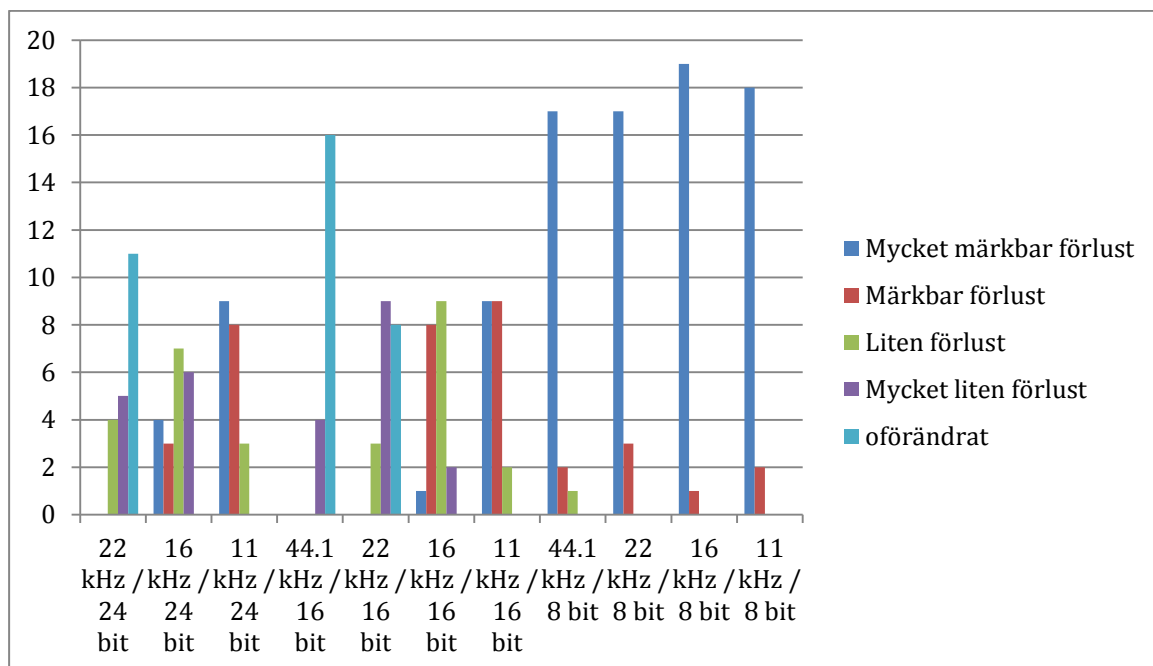
Tabell 4 Resultat för samtliga försökspersoner för det syntbaserade explosionsljudet.



5.2.3 Ljud: Inspelat effektljud

Liknande resultat som presenterats i de två tidigare ljuden återfinns även här. Grupp 1 bedömde ljud hårdare även här. Precis som de tidigare ljuden så återfinns ett liknande resultat, vid 24 bitdjup och hög samplingsfrekvens så är de flesta försökspersoner ense om att det antingen är oförändrat eller att det skett en liten eller mycket liten förlust i kvalitet. Även här syns ett liknande mönster för 16 bitdjup ljuden, vid en tillräckligt hög samplingsfrekvens märks det ytterst liten skillnad från referensljudet. Denna gång var båda grupper ense om att det på sin höjd hörs en mycket liten kvalitetsförlust vid 44.1 kHz med ett bitdjup på 16. Det observerades även att spridningen på svaren varierade mycket lite på just detta ljud. I de flesta fall användes bara 2-3 svarsalternativ av samtliga i de enskilda grupperna. Vid en sammanslagning av samtliga resultat är variationen dock högre då de båda grupperna bedömde kvalitetsförlusten olika. En viss skillnad i bedömningen av ljuden med 8 i bitdjup observerades denna gång. Vissa försökspersoner i grupp 2 bedömde att ljudkvalitén hade försämrats märkbart men inte lika mycket som de flesta andra. I grupp 1 gavs det ett liknande betyg, dock med färre åsikter som skiljde sig från mängden. Det bör dock poängteras att båda grupperna bedömde 8 bitdjup orsakade en mycket märkbar förlust i ljudkvalitet.

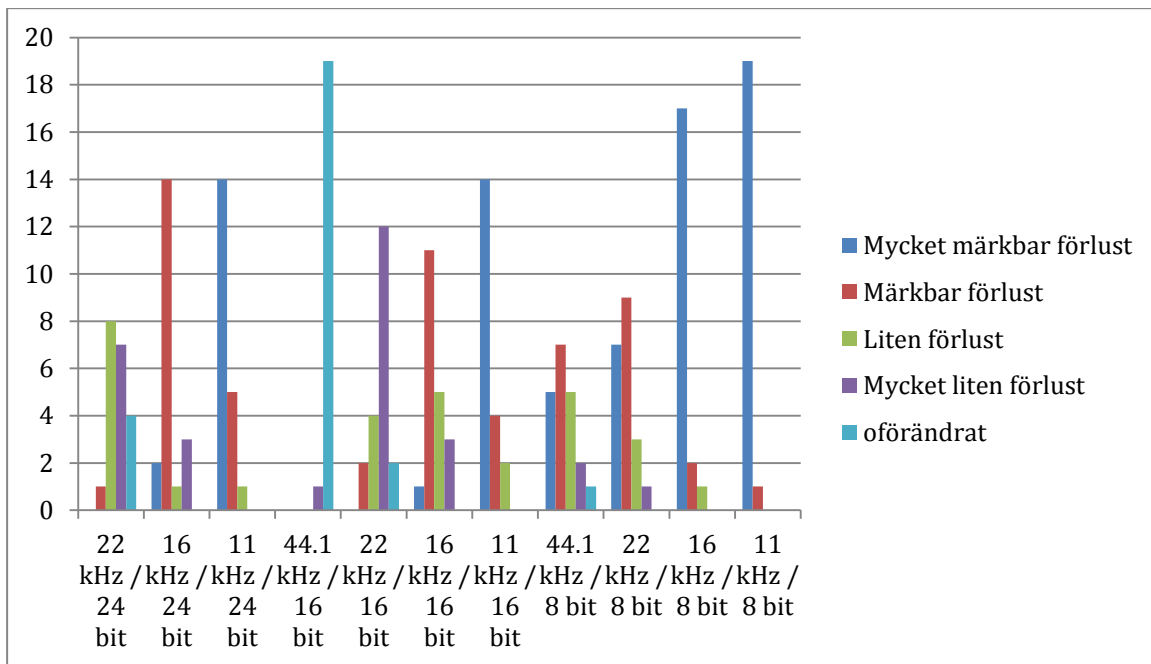
Tabell 5 Resultat för samtliga försökspersoner för det inspelade effektljudet.



5.2.4 Ljud: Syntbaserat effektljud

Även här ser vi samma mönster som i de tidigare ljud som presenterats, grupp 1 bedömde ljudet hårdare än grupp 2. Denna gång märktes det dock vissa skillnader. Båda grupperna var relativt likvärdiga i sin bedömning av ljuden med en minsta gräns på 16 bitdjup, därefter så började åsikterna gå isär. Det syntbaserade effektljudet skapade viss spridning i båda testgrupperna. Svårigheten syntes dock mest hos grupp 2 när de fick bedöma ljudkvalitén efter en kvalitetsförlust skett. Det yttrade sig som en stor spridning över de fem svarsalternativ som finns. Värt att notera är även att de poäng som lagts på olika alternativen är också relativt jämnt fördelade med vissa undantag. Förvånande nog är det även sant för ljuden med 8 bitdjup, även de ser en viss spridning vid en högre samplingsfrekvens. Det unika denna gång är förekomsten av spridning i resultatet för båda grupperna, i många fall finns det en eller ett fåtal bedömningar som sticker ut ifrån majoriteten. Denna gång finns det betydligt fler och de är mer utspridda än vad som syns på bedömningen av de andra ljuden. Det syns speciellt på de 8 bitars ljud med en hög samplingsfrekvens och blir progressivt färre ju sämre samplingsfrekvensen blir. Det tyder på att ljudet fungerar bättre vid ett lågt bitdjup än de andra ljuden, förutsatt att samplingsfrekvensen är tillräckligt hög. Vid en jämförelse med det syntbaserade explosionsljudet märks det ett visst mönster i spridningen vid de ljud där bitdjupet är lika med eller överstiger 16.

Tabell 6 Resultat för samtliga försökspersoner för det syntbaserade effektljudet.



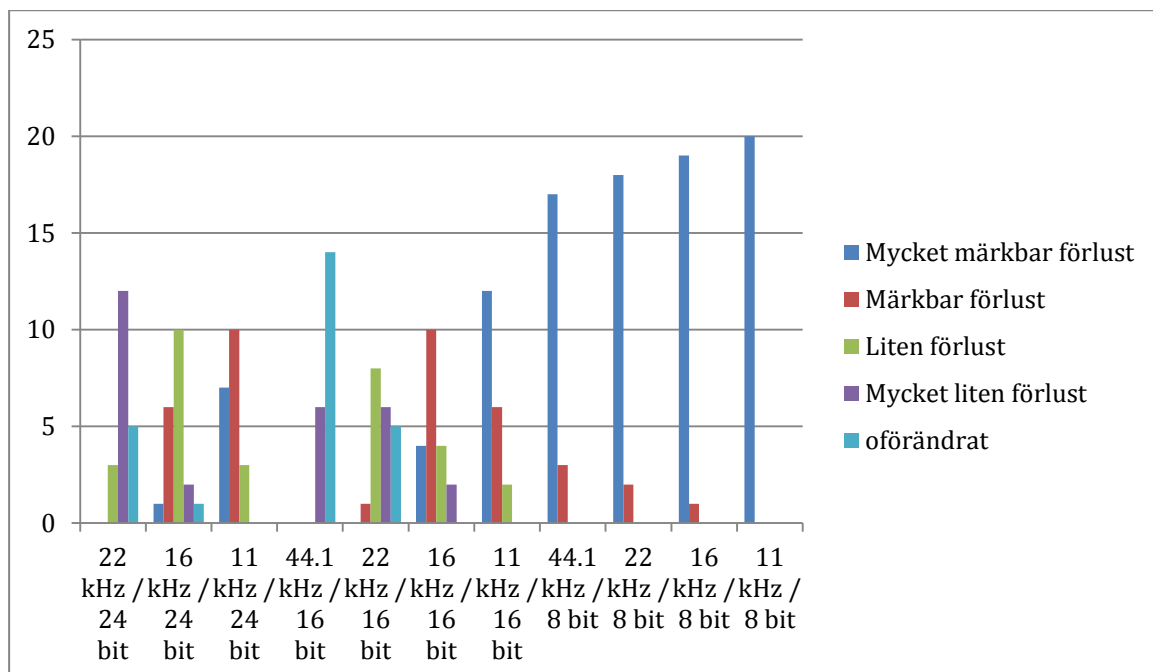
5.2.5 Ljud: Inspelat tal

Även denna gång bedömde grupp 1 kvalitetsförlusten hårdare än grupp 2. Speciellt synligt är det vid ljud med 8 bitdjup där samtliga i grupp 1 bedömde att de upplevt en mycket märkbar förändring. I övrigt var spridningen av bedömningen jämförbar med de tidigare ljuden, dock med få variationer i bedömningen i sig. Värt att notera är hur lika båda grupperna bedömde kvalitetsstegen 44.1 kHz, 22 kHz vid ett bitdjup av 16. Båda grupperna hade en nästintill identisk bedömning av båda kvalitetssteg, både i fördelning av poäng samt även i spridningen av dessa.

En av de viktigaste anledningarna till inkluderingen av det inspelade talet var undersökandet om huruvida ljud innehållande information, som mänskligt tal, påverkas i större grad av kvalitetsförlust. Vid granskandet av den data som samlats in tycks det inte ha påverkats mer än något annat ljud i undersökningen vid vissa kvalitetssteg. Speciellt vid 24 och 16 bitar är ljudet inte hårdare bedömt än något annat i undersökningen, i vissa fall till och med mindre hårt bedömt. Spridningen av de svar som försökspersonerna angett varierar inte heller avsevärt från något annat i undersökningen. Först vid 8 i bitdjup blev bedömningen hårdare än på de andra ljuden som användes till undersökningen. Speciellt synligt blev det när data från grupp 1 granskas, där samtliga ansåg att ljudet hade försämrats på ett mycket märkbart sätt.

Då ingen syntbaserad motsvarighet till det inspelade talet kunde skapas av en tillräckligt god kvalitet för undersökningen så uteblir den jämförelsen.

Tabell 7 Resultat för samtliga försökspersoner för det inspelade talet.



5.3 Summering

Summeringen av analysen är en sammanslagning av de analyser som skett tidigare för de enskilda ljud som användes i undersökningen. Då undersökningen samt arbetet är avsett för att kunna tillämpas i en riktig produktion bör ett flertal olika frågor analyseras. Huvudsakligen skall summeringen vara en jämförelse mellan ljudtyperna som syntbaserade ljud och inspelade ljud. I övrigt skall det även jämföras mellan de två grupperna av försökspersoner hur de bedömde ljud och vilken grupp som ansåg att kvalitén hade försämrats mest. Det skall även analyseras vilket eller vilka ljud som behöll mest kvalitet vid en förlust i den tekniska kvalitén och huruvida de har ett sammanband.

För att besvara en av de viktigaste frågorna som ställdes i början av undersökningen skall även en stor del av analysen ägnas åt att hitta en bra medelpunkt beträffande upplevd ljudkvalitet samt storlek på ljud. För att kunna implementeras i ett kommersiellt projekt med företaget Pyramid Beam Studios måste samtliga ljud analyseras för att se hur de påverkas och hur de kan användas till en mobil plattform. När en person eller företag jobbar mot en mobil plattform finns det ett flertal faktorer som måste tas i åtanke (Belcher, 2010). Det första är ljudkvalitén, speciellt anpassad för en mobil enhet. Då en mobilhögtalare inte kan spela upp ljud av en mycket hög kvalitet på ett trovärdigt sätt kan kvalitén sänkas till en viss gräns utan att det märks (Belcher, 2010). Att hitta den gräns där både kvalitet och en liten filstorlek möts är essensen i arbetet. Vad arbetet och resultatet av analysen inte kan ha hänsyn till är de skillnader som finns mellan olika mobila enheter och deras varierande förmåga att spela upp ljudfiler. Då förmågan kan variera kraftigt så tas inte den typen av variabler med i analysen utan får sedan anpassas beroende på enhet vid ett projekt. Det analysen skall tillföra är en så nära idealisk kombination av samplingsfrekvens och bitdjup som möjligt för en mobil plattform. Men som också samtidigt sparar på det begränsade utrymmet.

5.3.1 Jämförelse av ljudtyper

Vid en analys av de syntbaserade ljud som har använts i undersökningen så finns det en hel del intressanta resultat. Resultaten för de två syntbaserade ljud som användes skiljer sig från varandra på ett flertal sätt, framförallt det syntbaserade effektljudet urskiljer sig från mängden med en stor variation på bedömningen. Speciellt tydligt är det med de ljud som har ett bitdjup på 8. Enligt försökspersonerna så varierar bedömningen från oförändrad till en mycket märkbar förlust, något som är unikt för hela undersökningen just på ljud med 8 bitdjup. Besynnerligt nog bedömde försökspersonerna de kvalitetssteg av en högre kvalitet hade försämrats mer än vad de vanligtvis gjorde med de andra ljud som var med i undersökningen. Som ett exempel, vid 8 bitdjup och 44.1 kHz samt 22 kHz var det endast det syntbaserade effektljudet som inte hade ett tvåsiffrigt resultat på alternativet "Mycket märkbar förlust". Vad som är besynnerligt med just det resultatet är hur överens de två grupperna som helhet var över just kvalitetssteget 44.1 kHz och 16 bitdjup. Endast en försöksperson ansåg att kvalitetsförlusten var mycket liten medan resterande 19 personer ansåg den vara oförändrad. Om man jämför resultaten med det andra syntbaserade ljudet som användes i undersökningen så finns det inte många likheter förutom vissa delar av bedömningen. Vid 8 bitdjup märks den största skillnaden mellan de två syntbaserade ljuden, explosionsljudet ansågs ha förlorat mer kvalitet än effektljudet som hade klarat av förlusten av den tekniska kvalitén bäst av samtliga ljud. De båda har en gemensam nämnare i det att vid högre samplingsfrekvens så är bedömningen spridd över de olika svarsalternativen. Det tyder på att de syntbaserade ljuden antingen är svårare att bedöma eller att de klarar av en större förlust i teknisk kvalitet än andra typer av ljud. Vid en granskning av resultatet tycks

det visa att ljud av ett syntbaserat ursprung som använts vid undersökningen klarar av en förlust av bitdjup bättre än en förlust av samplingsfrekvens.

Vid en närmre observation syntes det att båda de inspelade ljuden, effektljudet och explosionen, bedömdes på ett mycket liknande sätt. Det förekom en mycket snarlik typ av spridning mellan åsikterna angående kvalitetsförlusten, även liknande mängd per svarsalternativ. En punkt där de skiljer sig från de syntbaserade ljuden är deras liknande bedömning över samtliga steg i bitdjup. Som tidigare nämnt var åsikterna spridda över flera svarsalternativ vid bedömningen av de syntbaserade ljuden, speciellt det syntbaserade effektljudet vid ett bitdjup av 8. Det bör också nämnas att de båda inspelade ljudtyperna bedömdes ha behållit mer kvalitet än de syntbaserade, dock började de bli mer och mer lika vid 8 bitdjup där kvalitetsförlusten är påtaglig. Detta är dock något som är sant för samtliga ljud med undantag för det syntbaserade effektljudet som klarade det lägre bitdjupet betydligt bättre än de andra enligt försökspersonerna.

De inspelade ljuden hade en mindre spridning i bedömningen av kvalitetsförlusten vid ett lägre bitdjup. Dock förekom det ett flertal svar som skiljde sig från majoriteten. De var dock få och närliggande till den bedömning som majoriteten gjort.

I jämförelsen av de båda typerna av ljud, inspelat och syntbaserat, var det tydligt att det syntbaserade effektljudet klarat förlusten av kvalitet på den allra lägsta nivån bäst. Samtidigt som ljudet klarat den lägsta nivån mycket bra i jämförelse med de andra ljuden som användes i undersökningen, så var det också ett av de ljud som bedömdes ha förlorat mest kvalitet vid de högre samplingsfrekvenserna och bitdjup. Den inspelade explosionen har fått en liknande bedömning av försökspersonerna vid hög samplingsfrekvens, vilket kan utsluta en koppling mellan att syntbaserat eller inspelat material skulle förlora mer upplevd ljudkvalitet i jämförelse med den andra. Anledningen till de varierande svaren beror troligtvis mer på de frekvenser som ljuden är uppbyggda av och hur de påverkas av en minskning i samplingsfrekvens eller bitdjup. Detta är något som ej togs med i undersökningen men som fanns i åtanke. Det valdes dock bort på grund av undersökningens tidramar men bör inkluderas i en möjlig framtida undersökning.

Det inspelade talet var unikt i det avseendet att det var i en ensam typ av grupp, ett ljud som innehåller information utöver det som kan tolkas på olika sätt beroende på vem som uppfattar ljudet. Det delar grupp med de övriga inspelade ljud som användes i undersökningen, men i övrigt så skiljer det sig i funktion från de andra ljuden. Det kan kopplas till IEZA-modellen Tol, Huiberts(2008), speciellt funktionen tal fyller i ett spelsammanhang. Tal fyller en viktig funktion i spel i flera avseenden vilket gör det mycket intressant huruvida det påverkas mer eller mindre av kvalitetsförlust. Vid en granskning och jämförelse med de övriga ljuden som användes i undersökningen tycks det inspelade talet inte ha påverkats annorlunda än de andra inspelade ljuden. Bedömningen följer samma mönster som för de övriga inspelade ljud som användes, försökspersonerna bedömde ljudet ofta likartat i deras respektive grupper. Extra tydligt är detta vid 8 bitdjup där samtliga i gruppen med erfarenhet av ljud bedömde att alla ljud hade en mycket märkbar kvalitetsförlust. Utöver de ljud som är längst ner på kvalitetsspektrumet i undersökningen så var den kombinerade gruppens bedömning mycket lika varandra. Huruvida det tyder på att ljud som tal, vilket de flesta människor hör varje dag, är lättare att bedöma än andra är otydligt. Den tydliga åsikten att ljudet får en mycket märkbar försämring vid lågt bitdjup kan tyda på en viss grad av extra känslighet när det kommer till mänskligt tal.

5.3.2 En medelpunkt

Den data som samlats in kan utöver utvärderingen av de två gruppernas uppfattningsförmåga kan också användas till att hitta en balans mellan ljudkvalitet och filstorlek. Belcher (2010) nämner att okomprimerade ljudformat som till exempel wav utgör ett problem i ljuddesign framförallt för mobila plattformar då ljudfiler av det formatet tenderar att ha mycket stora filstorlekar. En av de viktigaste aspekterna av arbetet var att hitta en den punkt där ljud och storlek ligger på en tillräcklig nivå för att kunna användas i en riktig spelproduktion. Ytterligare en aspekt att ha i åtanke är om olika ljud har olika medelpunkter vilket kan leda till att en eventuell produktion kan behöva ha olika kvalitet beroende på de olika ljud som kan behövas. Ett exempel kan vara om inspelat tal har en medelpunkt som kräver en högre kvalitet än ett ljud som endast fyller en funktion som ett effektljud. Om så är fallet kan det vara av företaget eller individens intresse att låta ljud som tal ta mer plats då det ofta fyller en mycket vital funktion i spel, som till exempel föra handlingen framåt eller informera spelaren om spelvärlden.

Vid en granskning av resultatet av undersökningen kan man tydligt se att det inte finns ett behov av ett mycket högt bitdjup. Vid 24 bitdjup är ljudet visserligen bra, speciellt vid 22 kHz samplingsfrekvens, och framförallt referensljudet vid 44.1 kHz och 24 bitdjup. Dock vid en analys av den data som samlats in ansåg en stor majoritet att ljudet var oförändrat vid 44.1 kHz och 16 bitdjup. Samtliga ljud bedömdes på ett mycket liknande sätt vid just detta kvalitetssteg, majoriteten ansåg att ljudet var helt oförändrat, samt en minoritet ansåg att det enbart var en mycket liten kvalitetsförlust. Övriga alternativ användes ej för bedömningen av det kvalitetssteget av försökspersonerna. Det syns tydligt på resultatet att det är just det kvalitetssteget där båda grupperna var i stort sett ense om kvalitetsförlusten där bitdjupet inte understiger 16. Uppenbarligen bör en produktion inte använda sig av ljud med ett bitdjup av 8 då resultatet visar att i stort sett alla försökspersoner anser att de har förlorat mycket kvalitet, oavsett vilken samplingsfrekvens man använder. Viss variation av bedömningen förekom givetvis, som också nämndes tidigare. Dock visar en överväldigande majoritet av försökspersonerna att ljud med ett så lågt bitdjup som 8 inte funkar.

Vart finns då punkten där kvalitet och filstorlek är på en tillräcklig nivå för en produktion till en mobil plattform? Som tidigare nämnt så är förutsättningarna för en ljudläggning till en mobil plattform mycket annorlunda i jämförelse med de plattformar som använder sig av en högtalare i fullstorlek och där filstorlek inte spelar lika stor roll. Detta bör också finnas i åtanke hos ljuddesignern som eventuellt vill använda sig av den data som tagits fram i detta arbete. Då en mobil plattform har en sådan begränsad förmåga att spela upp ljud på ett trovärdigt sätt kan en viss kvalitetsförlust döljas och på så sätt inte störa användaren (Belcher 2010). En bra startpunkt blir då det ljudet som de flesta försökspersoner var överens om att det höll god kvalitet, det vill säga 44.1 kHz och 16 bitdjup. Förutsatt att det är möjligt att dölja en viss förlust av kvalitet genom uppspelningen på en mobilhögtalare vore nästa steg att gå ner till 22 kHz och 16 bitdjup. Då frekvensåtergivningen och kvalitén blir betydligt sämre på en liten högtalare så blir skillnaden på ett ljud av referensljudets kvalitet och ljud med en kvalitet av 22 Khz och 16 bitdjup mycket små.

6 Slutsats

I denna del av arbetet presenteras resultatet av undersökningen och huruvida de frågeställningar som undersöktes har besvarats. Resultatets genomgång innefattar de frågeställningar som undersökningen använt för att hitta svar på. De besvaras i ordning och därefter huruvida någon ny kunskap uppnåts. Efter resultatet följer diskussionen om den kunskap som uppnåts med hjälp av undersökningen. Även hur den kan appliceras i ett verkligt sammanhang och vad kunskapen kan användas till utöver det som den användes till i undersökningssammanhanget. Slutligen tas möjliga framtida fortsättningar på arbetet och hur det kan expanderas och göras till något större.

6.1 Resultatsammanfattning

Undersökningens data visade på svar som i vissa fall var väntade och i andra fall mindre väntade. Båda grupperna av försökspersoner visade på ett mycket konsekvent bedömning av de ljudfiler som användes i undersökningen. Givetvis fanns det variationer i bedömningen, i vissa fall tycks det vara vissa ljud som var svårbedömda, i andra fall försökspersoner som bedömde olika ljud ovanligt hårt. Mest förvånade var den spridda bedömningen av det syntbaserade effektljudet som tycktes förvirra ett flertal försökspersoner i undersökningen. Ett flertal försökspersoner kommenterade att ljudet var klart och tydligt vid 8 bitdjup fram till det att efterklngen kom in där det uppstod ett långt sprakande. Det förvirrade en del av försökspersonerna till den grad att det trodde att det var ett helt annat ljud än det som spelats upp tidigare med ett bitdjup på 16 istället för 8. Intressant nog noterades det även kommentarer om att ljud med ett bitdjup av 8 ändå lät bra som ett effektljud oavsett om själva kvalitén hade försämrats. En del ansåg att ljudet till och med blivit bättre än originalet då det hade förändrats till ett helt nytt ljud. Det är något en ljuddesigner möjligtvis kan ha i åtanke när han eller hon skapar ljud, att förlusten av kvalitet kan användas till fördel för en ljudläggning av ett spel. Ett konkret exempel på detta; en sprakande radio med dålig mottagning. Om kvalitén på ljudet är menat att vara lågt kan det vara platsbesparande att faktiskt spara ner ljudet på ett sådant sätt att ljudet får en sådan karaktär naturligt.

Mindre förvånande var resultatet angående gruppernas bedömning och vilken grupp som bedömde olika kvalitetssteg olika hårt. Grupp 1 ansåg att kvalitetsförlusten märktes i större grad än grupp 2. Det kan kopplas till det resultat som Grabrielsson, Sjögren (1978) kom fram till med sin undersökning, där de såg att försökspersoner med en bakgrund i Hifi hade lättare för att upptäcka kvalitetsförlust i ljud.

Ytterligare resultat i undersökningen inkluderar huruvida vissa ljud klarar av en kraftig förlust i kvalitet baserat på om de skapats med inspelat ljudmaterial eller genererats med hjälp av en mjukvarusynt. Resultatet var otydligt då de ljud som tillhör de olika grupperna varierade i bedömningen. Det ljud som stod ut från de andra var det syntbaserade effektljudet som klarade låga nivåer av bitdjup betydligt bättre än något annat ljud i undersökningen. Det andra syntbaserade ljudet klarade dock inte av låga bitdjup lika väl, vilket talar emot att just syntbaserade ljud skulle klara av det bättre än inspelade ljud. Anledningen till hur just det syntbaserade effektljudet klarade av det låga bitdjupet är oklart och bör undersökas vidare vid en eventuell expansion eller fortsättning av arbetet.

En del av arbetet som är kopplat till föregående punkt berör vilken nivå och teknisk del av ljudet som är viktigast för en god ljudkvalitet. Som tidigare nämnt ansåg en stor majoritet att ljud vid 8 bitdjup hade utsatts för en sådan kvalitetsförlust att det var mycket märkbart i nästintill alla fall. Den data undersökningen genererat gör det lätt att utesluta ljud med just 8 bitdjup då kvalitetsförlusten är för utmärkande. Vid ett bitdjup av 16 däremot syns ett helt annat resultat, vid en tillräckligt hög samplingsfrekvens ansåg en majoritet av försökspersonerna att ljudet var oförändrat i kvalitén. Med den kunskapen kan vi även utesluta ljud med ett bitdjup av 24 då det inte är nödvändigt med ett så högt bitantal för ljudkvalitén, speciellt inte vid en produktion för en mobil enhet där skillnaden mellan 16 och 24 bitdjup är ännu svårare att urskilja. Nästa aspekt angående kvalitén är samplingsfrekvensen och var den bör ligga för att få ett så kvalitativt ljud som möjligt. Vid 44.1 kHz som tidigare nämnt ansåg majoriteten att ljudet var oförändrat, och vid 22 kHz var bedömningen mer utspridd, dock ansågs kvalitén inte ha försämrats så den var mycket märkbar. Allt under 22 kHz ansågs vara märkbart och reflekteras i den framtagna datan. Slutsatsen av det är för att få ett ljud av tillräckligt god kvalitet krävs det en tillräckligt hög nivå av både bitdjup och samplingsfrekvens. Den ena utan den andra väger väldigt sällan upp ljudet så det får en tillräckligt hög kvalitet för en spelproduktion. Nyckeln till ett ljud med god ljudkvalitet är en kombination av båda i tillräckligt höga nivåer. Detta reflekteras i den kvalitet som används för Cd-skivor, 44.1 kHz med 16 bitdjup. Vid en produktion för en mobil enhet bör dock en ljuddesigner tänka på att ljud med 16 bitdjup och 22 kHz är något sämre, men också mindre i storlek. En avvägning bör göras beroende på hur många ljud som skall användas, samt hur stort fokus ljudet har i spelet för att bedöma vilken nivå som skall användas.

6.2 Diskussion

Arbetet är huvudsakligen menat att göra nytta och ge ljuddesigners ett redskap för att avväga vilken del av ett ljud de kan skära ner på för att minska storlek samt bibehålla god ljudkvalitet. Det ska även fungera som ett redskap för att avväga om ett ljud bör vara av ett speciellt ursprung, som till exempel syntbaserat eller inspelat. Också vilket av dessa som är mest lämpade för olika funktioner som effekter eller interfaceljud.

Förhoppningen är att det skall användas som ett redskap eller riktlinje för riktiga produktioner av spel, framförallt mobilspel där avvägningen mellan ljudkvalitet och filstorlek är av yttersta vikt. Samtidigt vara tillräckligt förståeligt så att både vana ljuddesigners och nybörjare skall kunna dra nytta av arbetet och på så sätt utvecklas inom området.

I hans artikel tar Belcher (2010) upp många viktiga aspekter av ljuddesignen för en mobil plattform. Det som han inte tar upp är ett undersökningsresultat med försökspersoner angående vad det är som gör att vi upplever att ljudkvalitet försämras och i vilken grad. Förhoppningsvis kan detta arbete komplettera det Belcher beskriver med sin personliga erfarenhet av ljuddesign med data som samlats in med försökspersoner för att ge en bättre bild över hur människor upplever ljudkvalitet.

6.3 Framtida arbete

Det som undersökningen hade tjänat mest på vid en framtida utveckling eller fortsättning är mer tid. Det hade lett till ett flertal möjliga förbättringar för undersökningen. I ett tidigare kapitel visades hur många ljud som blev bortvalda i mån av tid, en uppenbar åtgärd vore att inkludera de ljud som sållades bort och därmed att få en mer grundlig undersökning. Ytterligare kvalitetssteg hade också varit något som gjort undersökningen mer tillförlitlig samt gjort medelpunkten för kvalitet och filstorlek mer exakt.

Ytterligare förbättringar som kan göra undersökningen mer tillförlitlig är en större mängd försökspersoner av en mer varierad bakgrund. Möjligen skulle undersökningen kunna inkludera de med hörselskador som tinnitus för att se om de reagerar annorlunda på kvalitetsförlust och om deras medelpunkt är på en annan plats än för de utan hörselskador. Även en genus aspekt skulle kunna appliceras på undersökningen för att se om någon utav könen visar större känslighet för kvalitetsförlust.

En aspekt som kunnat bredda arbetet är en frekvensanalys av de ljud som använts i undersökningen för att se om det har något med olika bedömningar av ljudkvalitén att göra. Olika ljud från olika ursprung men med liknande frekvensinnehåll kan möjligtvis ge ett liknande resultat vilket kan ge ett möjligt sammanband mellan vissa ljud och varför de klarar av kvalitetsförlust på olika sätt.

Arbetet kan och kommer implementeras vid arbetet på Pyramid Beam Studios där fokus för företaget ligger på att skapa kvalitativa mobilspel. Arbetets avsikt har från början varit att användas i ett riktigt spelprojekt som en mall eller riktlinjer för vilken kvalitet ljud bör sparas i. Någon fortsättning på undersökningen i sig kommer inte ske inom en överskådlig framtid, dock kommer resultatet av den existerande undersökningen användas till företagets första och andra spel.

Referenser

Belcher, P.J. (2010). *iOS Audio Design: What Everyone Needs To Know*. Tillgänglig på internet:http://www.gamasutra.com/view/feature/6226/ios_audio_design_what_everyone_.php

Blauert, J. (1996) *Spatial hearing: the psychophysics of human sound localization*. Massachusetts: MIT Press.

Bryman, A. (2001) *Samhällsvetenskapliga metoder*, Malmö: Liber.

Cubase 5 (2009) [Datorprogram] Steinberg GmbH.

Gabrielsson, A, Sjögren, H. (1978) *Perceived sound quality of sound-reproducing systems*. Melville: Journal of the acoustical society of america, Volume 65, issue 4.

Tillgänglig på internet: <http://h24-files.s3.amazonaws.com/42685/100195-KhwRJ.pdf>

Rice, S. (2005), *A survey course on computer audio*, Mississippi: Consortium for Computing Sciences in Colleges , USA

Tol, R., Huiberts, S. (2008) *IEZA: A framework for game audio*. Tillgänglig på internet: http://www.gamasutra.com/view/feature/131915/ieza_a_framework_for_game_audio.php

Rubin, D. & Tully, T. (1996), *The audible PC*, Berkley: Sybex Inc.

Propellerhead Reason 5 (2010) [Datorprogram] Propellerhead Software.

Helland, K., Knapskog, K., Larsen, L. O. & Østbye, H. (2003) *Metodbok för medievetenskap*. Malmö: Liber

Wikipedia - Quantization

http://en.wikipedia.org/wiki/Quantization_%28signal_processing%29

[Figur 1: Hämtad 2012-05-29]

[Figur 2: Hämtad 2012-05-29]

Appendix A - Undersöknings resultat

Grupp utan erfarenhet						
Inspelad explosion						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit			4	3	3
2	16 kHz / 24 bit		3	1	4	2
3	11 kHz / 24 bit	5	1	4		
4	44.1 kHz / 16 bit	1			3	6
5	22 kHz / 16 bit				5	5
6	16 kHz / 16 bit	3	2	1	4	
7	11 kHz / 16 bit	6	2	1	1	
8	44.1 kHz / 8 bit	8	2			
9	22 kHz / 8 bit	9	1			
10	16 kHz / 8 bit	10				
11	11 kHz / 8 bit	10				
Synt explosion						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit		1	3	2	4
2	16 kHz / 24 bit		4	3	2	1
3	11 kHz / 24 bit	6	2	1	1	
4	44.1 kHz / 16 bit				4	6
5	22 kHz / 16 bit		1	3	4	2
6	16 kHz / 16 bit	2	2	4	1	1
7	11 kHz / 16 bit	4	4	1	1	
8	44.1 kHz / 8 bit	10				
9	22 kHz / 8 bit	9	1			

10	16 kHz / 8 bit	8	2			
11	11 kHz / 8 bit	9		1		
Inspelad effekt						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit				3	7
2	16 kHz / 24 bit		2	3	5	
3	11 kHz / 24 bit	5	2	3		
4	44.1 kHz / 16 bit				2	8
5	22 kHz / 16 bit				5	5
6	16 kHz / 16 bit		3	6	1	
7	11 kHz / 16 bit	3	5	2		
8	44.1 kHz / 8 bit	8	2			
9	22 kHz / 8 bit	8	2			
10	16 kHz / 8 bit	9	1			
11	11 kHz / 8 bit	9	1			
Synt effekt						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit			4	2	4
2	16 kHz / 24 bit		7		3	
3	11 kHz / 24 bit	8	1	1		
4	44.1 kHz / 16 bit				1	9
5	22 kHz / 16 bit		1	2	6	1
6	16 kHz / 16 bit		6	2	2	
7	11 kHz / 16 bit	8	1	1		
8	44.1 kHz / 8 bit	1	5	2	2	
9	22 kHz / 8 bit	2	6	1	1	
10	16 kHz / 8 bit	9		1		

11	11 kHz / 8 bit	9	1			
Inspelat tal						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit				7	3
2	16 kHz / 24 bit	1		7	1	1
3	11 kHz / 24 bit	1	7	2		
4	44.1 kHz / 16 bit				3	7
5	22 kHz / 16 bit			4	3	3
6	16 kHz / 16 bit	1	5	3	1	
7	11 kHz / 16 bit	5	4	1		
8	44.1 kHz / 8 bit	7	3			
9	22 kHz / 8 bit	8	2			
10	16 kHz / 8 bit	9	1			
11	11 kHz / 8 bit	10				

Grupp med erfarenhet						
Inspelad explosion						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit		1	3	5	1
2	16 kHz / 24 bit	3	5	2		0
3	11 kHz / 24 bit	6	4			
4	44.1 kHz / 16 bit				4	6
5	22 kHz / 16 bit			2	3	5
6	16 kHz / 16 bit		7	3		
7	11 kHz / 16 bit	8	2			
8	44.1 kHz / 8 bit	10				
9	22 kHz / 8 bit	10				

10	16 kHz / 8 bit	10				
11	11 kHz / 8 bit	9	1			
Synt explosion						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit		1	1	3	5
2	16 kHz / 24 bit	2	4	2	2	
3	11 kHz / 24 bit	5	5			
4	44.1 kHz / 16 bit				4	6
5	22 kHz / 16 bit		4	1	3	2
6	16 kHz / 16 bit	3	1	4	2	
7	11 kHz / 16 bit	6	4			
8	44.1 kHz / 8 bit	10				
9	22 kHz / 8 bit	10				
10	16 kHz / 8 bit	9	1			
11	11 kHz / 8 bit	9	1			
Inspelad effekt						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit			4	2	4
2	16 kHz / 24 bit	4	1	4	1	
3	11 kHz / 24 bit	4	6			
4	44.1 kHz / 16 bit				2	8
5	22 kHz / 16 bit			3	4	3
6	16 kHz / 16 bit	1	5	3	1	
7	11 kHz / 16 bit	6	4			
8	44.1 kHz / 8 bit	9		1		
9	22 kHz / 8 bit	9	1			
10	16 kHz / 8 bit	10				

11	11 kHz / 8 bit	9	1			
Synt effekt						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit		1	4	5	
2	16 kHz / 24 bit	2	7	1		
3	11 kHz / 24 bit	6	4			
4	44.1 kHz / 16 bit					10
5	22 kHz / 16 bit		1	2	6	1
6	16 kHz / 16 bit	1	5	3	1	
7	11 kHz / 16 bit	6	3	1		
8	44.1 kHz / 8 bit	4	2	3		1
9	22 kHz / 8 bit	5	3	2		
10	16 kHz / 8 bit	8	2			
11	11 kHz / 8 bit	10				
Inspelat tal						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit			3	5	2
2	16 kHz / 24 bit		6	3	1	
3	11 kHz / 24 bit	6	3	1		
4	44.1 kHz / 16 bit				3	7
5	22 kHz / 16 bit		1	4	3	2
6	16 kHz / 16 bit	3	5	1	1	
7	11 kHz / 16 bit	7	2	1		
8	44.1 kHz / 8 bit	10				
9	22 kHz / 8 bit	10				
10	16 kHz / 8 bit	10				
11	11 kHz / 8 bit	10				

Samtliga försökspersoner						
Inspelad explosion						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit		1	7	8	4
2	16 kHz / 24 bit	3	8	3	4	2
3	11 kHz / 24 bit	11	5	4		
4	44.1 kHz / 16 bit	1			7	12
5	22 kHz / 16 bit			2	8	10
6	16 kHz / 16 bit	3	9	4	4	
7	11 kHz / 16 bit	14	4	1	1	
8	44.1 kHz / 8 bit	18	2			
9	22 kHz / 8 bit	19	1			
10	16 kHz / 8 bit	20				
11	11 kHz / 8 bit	19	1			
Synt explosion						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit		2	4	5	9
2	16 kHz / 24 bit	2	8	5	4	1
3	11 kHz / 24 bit	11	7	1	1	
4	44.1 kHz / 16 bit				8	12
5	22 kHz / 16 bit		5	4	7	4
6	16 kHz / 16 bit	5	3	8	3	1
7	11 kHz / 16 bit	10	8	1	1	
8	44.1 kHz / 8 bit	20				
9	22 kHz / 8 bit	19	1			
10	16 kHz / 8 bit	17	3			
11	11 kHz / 8 bit	18	1	1		

Inspelad effekt						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit			4	5	11
2	16 kHz / 24 bit	4	3	7	6	
3	11 kHz / 24 bit	9	8	3		
4	44.1 kHz / 16 bit				4	16
5	22 kHz / 16 bit			3	9	8
6	16 kHz / 16 bit	1	8	9	2	
7	11 kHz / 16 bit	9	9	2		
8	44.1 kHz / 8 bit	17	2	1		
9	22 kHz / 8 bit	17	3			
10	16 kHz / 8 bit	19	1			
11	11 kHz / 8 bit	18	2			
Synt effekt						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit		1	8	7	4
2	16 kHz / 24 bit	2	14	1	3	
3	11 kHz / 24 bit	14	5	1		
4	44.1 kHz / 16 bit				1	19
5	22 kHz / 16 bit		2	4	12	2
6	16 kHz / 16 bit	1	11	5	3	
7	11 kHz / 16 bit	14	4	2		
8	44.1 kHz / 8 bit	5	7	5	2	1
9	22 kHz / 8 bit	7	9	3	1	
10	16 kHz / 8 bit	17	2	1		
11	11 kHz / 8 bit	19	1			

Inspelat tal						
		Mycket märkbar förlust	Märkbar förlust	Liten förlust	Mycket liten förlust	Oförändrad
1	22 kHz / 24 bit			3	12	5
2	16 kHz / 24 bit	1	6	10	2	1
3	11 kHz / 24 bit	7	10	3		
4	44.1 kHz / 16 bit				6	14
5	22 kHz / 16 bit		1	8	6	5
6	16 kHz / 16 bit	4	10	4	2	
7	11 kHz / 16 bit	12	6	2		
8	44.1 kHz / 8 bit	17	3			
9	22 kHz / 8 bit	18	2			
10	16 kHz / 8 bit	19	1			
11	11 kHz / 8 bit	20				