

MP3

Hur låter det?
How does it sound?



10-poängsuppsats informatik
2002-01-14

Författare

Rikard Delbring & Lars Johansson

Handledare

Christian Östlund

Examinator

Kerstin Grundén

Abstract

Mp3 is a compressionformat that gets more users. More and more downloads music from the Internet and the format is also used in presentations and i many other contexts. In our essay we have examined if people experience a difference between the original and the compression. We have also examined if there are divergence between speech and different categories of music and if the system on which the sound is played has importance. To get the answers for this we have carried out a experiment where thirty participants, in two different categories, *Students* and *Soundprofessionals*, listen to original and compressed sound in three different categories to see if they experience any divergence. Soundfiles in the categories speech, popular music and classical music were played on two different soundsystems to examine the importance of the system. The result of the experiment shows that mp3-compression works well in many ways. The experiment showed that speech worked well on both systems when few of the participants did not experience that the original sounded better. The classical music showed similar result with the exception of 96 kbps compression on the more advanced speakers where a vast majority thought that the original sounded better. Popular music was the category that the participants felt the original sounded better to the highest degree, but even there it works well at some compressions.

Sammanfattning

Mp3 är ett komprimeringsformat som får allt fler användare. Vår uppsats bygger på ett experiment där vi undersökt om folk upplever en skillnad mellan original och komprimerat ljud Vi har också undersökt om det är skillnad på upplevelsen mellan tal och olika musikformer och om uppspelningssystemet har betydelse. Ljudfiler i kategorierna tal, populärmusik och klassisk musik spelades upp på två olika ljudsystem. Resultatet av experimentet visar att mp3-komprimering fungerar bra i många sammanhang. Experimentet visade att tal fungerade bra på båda systemen då få av respondenterna inte upplevde att originalet lät bättre. Den klassiska musiken visade liknande resultat med undantag av 96 kbps komprimering på de avancerade högtalarna där en klar majoritet uppfattade originalet bättre. Den kategori som respondenterna i störst grad upplevde originalet som bäst var i populärmusik, men även där fungerar mp3-komprimering bra vid vissa komprimeringsgrader

Innehållsförteckning

1.	INTRODUKTION	1
2.	SYFTE & PROBLEMFÖRMULERING	1
2.1.	Uppsatsens målgrupp	1
2.2.	Problemformulering	1
2.3.	Avgränsning	2
3.	METOD & STRATEGI	2
3.1.	Strategi	2
3.2.	Metod	2
3.3.	Tillvägagångssätt	3
3.4.	Undersökningens uppläggning	3
3.4.1.	Datainsamling	3
3.4.2.	Urval	3
3.4.3.	Testgrupp	4
3.4.4.	Tillvägagångssätt	4
3.5.	Utformning av experiment	4
3.5.1.	Val av komprimeringsgrader och encoder	4
3.5.2.	Val av ljudfiler	5
3.5.3.	Val av uppspelningsordning	5
3.5.4.	Val av uppspelningsutrustning	6
3.5.5.	Experimentplats	6
3.5.6.	Förändringar efter testexperiment	7
3.6.	Databehandling	7
3.6.1.	Svarsbearbetning	7
3.7.	Kvalitetsgranskning	7
4.	TEORI	8
4.1.	Begreppsdefinitioner	8
4.1.1.	Ljud	8
4.1.2.	Frekvens	8
4.1.3.	Amplitud & Decibel	8
4.1.4.	Audio	8
4.1.5.	Psykoakustik	8
4.2.	Mänskligt hörande	9
4.2.1.	Frekvensområden	10
4.3.	Digitaliserat ljud	10
4.3.1.	Samplat ljud	10
4.3.2.	Samplingsfrekvens	10
4.3.3.	Bitdjup & upplösning	10
4.4.	Komprimering	11
4.4.1.	MPEG	11
4.5.	Mp3	12
4.5.1.	Bithastighet	12
4.5.2.	Filstorlek	13
4.5.3.	Ljudkvalitet	13
5.	RESULTAT	13
5.1.	Åldersgrupper	14
5.2.	Tal på PC högtalare	14
5.3.	Populärmusik på PC högtalare	14
5.4.	Klassisk musik på PC högtalare	15
5.5.	Tal på studiomonitorer	15
5.6.	Populärmusik på studiomonitorer	15
5.7.	Klassisk musik på studiomonitorer	16
6.	RESULTATDISKUSSION	16
6.1.	Metoddiskussion	16
6.1.1.	Respondenter	16
6.1.2.	Experimentutrustning och plats	17

6.1.3.	Utförande av experiment	17
6.2.	Åldersgrupper.....	17
6.3.	Respondenter.....	17
6.4.	Tal på Pc:högtalare.....	17
6.5.	Populärmusik på Pc:högtalare	18
6.6.	Klassisk musik på Pc:högtalare.....	18
6.7.	Tal på studiomonitorer.....	19
6.8.	Populärmusik på studiomonitorer	19
6.9.	Klassisk musik på studiomonitorer	20
6.10.	Komprimerat tal	20
6.11.	Komprimerad populärmusik.....	21
6.12.	Komprimerad klassisk musik	22
7.	SLUTSATS.....	23
8.	REFERENSER.....	25

1. Introduktion

Ljudet betyder mycket för det allmänna intrycket. Hörseln är ett av våra viktigaste sinnen och det ger oss både faktainformation och känsloupplevelse. Inom underhållningsindustrin har man länge varit medveten om att ljudet kan förhöja den visuella upplevelsen. I dagens multimedietitlar används musik och ljudeffekter för att öka dramatiken och realismen i bilderna, precis som på film (Kindersley, 1997).

Multimedia är en kombination av olika media som text, ljud och eller rörliga video sekvenser. En viktig ingrediens i en multimedieproduktion är ljudet (musik, speaker, ljudeffekter). Rätt musik och ljudeffekter kan avsevärt höja kvaliteten och intresset för en produktion. Musik har en fantastisk förmåga att försätta oss i olika sinnestämningar och själva stämningen byggs till stor del upp med ljud. (Bodelius, 1998)

Komprimering av ljuddata görs för att man vill minska överföringstid och spara plats på lagringsmedia. Själva komprimeringen kan göras på två olika sätt, förlustfri och förlustgivande. Vid komprimering av ljud medför detta oftast en kompromiss mellan kvalitet och prestanda. (Nilsson, 2000).

Mp3-komprimering, som är en förlustgivande komprimering är det i särklass vanligaste sättet att komprimera ljud och används inom en rad områden såsom animationer, videoproduktioner, e-learning, distribution av musik och ljud på Internet och olika former av presentationer. Om man vill använda sig av ljud i sin produktion är det därför viktigt att man är säker på att återgivningen hos användaren stämmer överens med det ljud som man hade när produktionen gjordes. (Hacker, 2000)

2. Syfte & Problemformulering

Examensarbetets syfte är att dels utreda hur olika grader av Mp3-komprimering upplevs på: tal och musik som kan användas för e-learning, multimedia-presentationer samt distribution på Internet. Det andra syftet med uppsatsen är att visa vilka komprimeringsgrader man kan använda sig av i sin produktion för att förmedla ett rättvisande ljud för sin tänkta målgrupp. Detta gör att resultatet av uppsatsen kan användas som ett riktmärke för val av komprimeringsgrad vid olika ljudproduktioner.

2.1. Uppsatsens målgrupp

Målgruppen för denna uppsats är de som arbetar med framställning av multimedieproduktioner, ljuddistribution på Internet samt olika former av e-learning projekt.

2.2. Problemformulering

Denna uppsats skall undersöka och belysa följande huvudfrågeställning:

- Vid vilka grader av Mp3-kompressioner upplever lyssnaren en försämring gentemot originalet på tal, populärmusik och klassisk musik?

2.3. Avgränsning

För att denna studie skall vara praktiskt genomförbar och leda fram till ett tolkbart resultat har följande avgränsningar gjorts:

- Uppsatsen kommer inte att utreda hur olika högtalarsystem låter i jämförelse med varandra och ej heller hur olika ljudkort hanterar uppspelning av komprimerat ljud.
- Uppsatsen skall inte analysera tekniken bakom MP3.
- Uppsatsen kommer inte att ha den omfattning att generella slutsatser om hur olika komprimeringsgrader låter kan slås fast.

3. Metod & Strategi

3.1. Strategi

Forskningsstrategin i detta arbete är av experimentell karaktär där tanken är att studera enskilda faktorer och dess effekter. Inom experimentell forskning utgår man ifrån att det finns en skillnad mellan individens upplevelse av omvärlden och den faktiska verkligheten. Denna skillnad är verifierbar via vetenskapliga metoder. Möjligheten att få objektiv kunskap om verkligheten sker genom observationer och mätningar som är objektiva. Observationer och sensoriska data, definierade som information förvärvad via våra sinnesorgan är det enda sättet att nå sann kunskap om verkligheten. Genom att använda accepterade och standardiserade forskningstekniker kan forskaren eliminera bias (uppfattningar som saknar vetenskapligt stöd) och på så vis komma fram till resultat via objektiva, kvantitativa mätningar. (DePoy & Gitlin, 1999)

Det ideala experimentet bör äga rum under så kontrollerade förhållanden som möjligt. Detta kan utföras i t ex ett avskilt rum där inga störningar inträffar; vilket brukar kallas laboratorieexperiment. (Patel & Tebelius, 1987)

I och med att varaktigheten på experimentet är relativt kort samt kommer att genomföras på en bestämd plats med hård kontroll över variabler för att isolera orsaksfaktorer. (Denscombe, 2000)

3.2. Metod

Vid val av metod har vi tagit i beaktande att välja en metod som är relevant för den frågeställning som uppsatsen har samt att den skall ge möjlighet till att samla in relevanta data. I och med att arbetet är tänkt att analysera olika personers uppfattning om hur olika komprimeringar låter kommer experimentet att bygga på okomplicerade svar med identiska frågor till alla respondenter. Detta gör att användning av ett frågeformulär med enbart fasta frågor är lämpligt. (Denscombe, 2000)

Vi kommer att använda oss av en kvantitativ forskningsmetod i och med att resultatet kommer att ha följande struktur:

- Resultatet av forskningen kan analyseras med hjälp av siffror
- Forskningen kommer att fokusera på specifika faktorer och kommer att studeras i förhållande till andra specifika faktorer.
- Resultatet av forskningen kommer att vara objektivt och inte avspegla vår egna uppfattning.

- Forskningen kommer att ha en förutbestämd design. (Denscombe, 2000)

Resultatet av detta arbete är tänkt att kunna forma en teori om hur man kan använda sig av ljudkomprimering i olika produktioner vilket gör att man kan säga att detta arbete har en induktiv ansats. (Patel & Davidson, 1994)

3.3. Tillvägagångssätt

Att mäta människors upplevelse är komplicerat och kräver noggranna förberedelser. Det som är viktigt är att man som forskare inte påverkar resultatet med sitt tillvägagångssätt. (DePoy & Gitlin, 1999)

Utifrån detta har vi diskuterat fram följande faktorer som måste tas i beaktning för vårt experiment:

- Val av musikstycke
- Val av talstycke
- Val av lokal för experiment
- Val av utrustning för uppspelning av ljudinformationen
- Antal respondenter vid varje experimenttillfälle
- Utformande av svarsformulär
- Ordningföljden för mätljuden
- Tidsåtgång för varje experimenttillfälle

3.4. Undersökningens uppläggning

För att reda ut begrepp och få en förståelse har vi studerat sekundärdata i form av lämplig litteratur och tagit del av artiklar som har haft anknytning till vårt undersökningsområde. Detta för att få inspiration och se hur andra författare har gått tillväga.

3.4.1. *Datainsamling*

I vår undersökning ligger tyngdpunkten på den primärdata som samlas in genom studie av verkligheten. Det är den empirin som ligger till grund i resultat och diskussions avsnitten.

3.4.2. *Urval*

Respondenturvalet kommer att vara baserat på syftet för uppsatsen. Vi kommer att göra ett subjektivt första urval av individer som passar för undersökningens syfte (Denscombe, 2000).

För att vi inte skall påverka urvalsgruppen i all för stor utsträckning kommer vi att genom snöbollsurval låta våra respondenter ge namn på ytterligare respondenter som uppfyller kriterierna för uppsatsens syfte (DePoy & Gitlin, 1999).

Respondenterna kan vara användare inom e-learning, multimediepresentationer samt musik distribuerat på nätet. Valet av dessa respondenter grundas på att de kommer i kontakt med Mp-3 komprimerat ljud. För att få fram ett kvantitativt resultat kommer vi att ha ca 10-15 respondenter från varje urvalsgrupp.

Den ena urvalsgruppen kommer att bestå av personer som i sitt arbete eller på en stor del av sin fritid kommer i kontakt med musik. Vi anser att dessa personer med stor erfarenhet av musik skall ha goda möjligheter att svara på uppsatsens frågeställning när man kan skilja en försämring av ljudet.

Den andra urvalsgruppen kommer att bestå av personer som kommer i kontakt med ljud i undervisningssyfte i sin skol- eller arbetsmiljö. Vi menar att de med sin kontakt med ljud i form av tal skall kunna avgöra när det sker en nedgång i kvalitet.

3.4.3. Testgrupp

Vid utformning av enkäten är det en god idé att pröva frågorna på en utvald testgrupp. Syftet är att kontrollera att respondenterna uppfattar frågorna på det sätt det är tänkt och att vi får de svar som enkäten avser att mäta. Detta moment genomförs för att få högre validitet i undersökningen. Det bör eftersträvas att personerna i testgruppen är jämförbara med den verkliga gruppen som skall undersökas. (Ejlertsson, 1996)

3.4.4. Tillvägagångssätt

Vid varje undersökningstillfälle kommer vi att ha max fem respondenter för att på det sättet kunna ha kvantitet per tillfälle utan att för den skull påverka respondenternas ljudupplevelse. Respondenterna kommer att sitta bredvid varandra, men något avskärmade så att de inte kan se varandras svar. Innan undersökningen börjar kommer vi poängtera för respondenterna att det inte kommer att redovisas några enskilda svarsresultat för att de inte skall känna någon press på att svara rätt, utan istället fokuserar på sina intryck av experimentet.

Eftersom vi vill undersöka upplevelsen av komprimerat ljud finns det egentligen inget givet, rätt svar. Vi vill med undersökningen få in olika individers subjektiva intryck av komprimerat ljud i jämförelse mot originalljudet, för att därigenom kunna skapa oss en uppfattning om vid vilken komprimeringsgrad respondenterna kan höra skillnad på olika ljud i olika system.

3.5. Utformning av experiment

Undersökningen kommer att gå till på det viset att vi kommer att spela upp ett originalljud som kan ses som ett referensljud för undersökningen. Sedan kommer vi slumpmässigt att spela upp samma ljudsnutt i olika komprimeringsgrader i par med originalljudet. Varje ljudfil kommer att vara ca 30 sek lång. Respondenterna ska sedan svara om de hör någon skillnad mellan de två ljudsnuttarna och om så är fallet notera vilket ljud som upplevs bäst. För att öka reliabiliteten i vårt experiment kommer vi också ha med ett ljudpar där vi jämför originalljudet mot sig självt för att se om respondenterna tycker sig höra någon skillnad originalljuden emellan samt att detta medför att respondenterna inte vet att det alltid kommer att finnas ett komprimerat ljud i varje uppspelat par.

För att särskilja de två olika systemen som kommer att användas för uppspelning av ljudfilerna kommer vi att på det första systemet använda oss av datorns ljudkort och högtalare. På det andra systemet kommer vi att bränna ut de olika ljudspåren på en CD skiva och spela upp dessa via en vanlig CD spelare.

3.5.1. Val av komprimeringsgrader och encoder

Experimentet kommer att utföras med olika komprimeringsgrader. Valet av dessa grundar sig på en undersökning som vi gjort på Internet via Audiogalaxy. Vi har provat att söka på de fem mest spelade låtarna på P3 under år 2000 och sett vilka komprimeringar som fanns tillgängliga, vilket var följande: 64, 96, 128, 160, 192, 256 och 320 kbps. Vi ansåg att sju komprimeringar tillsammans med originalljudparet blev för mycket, så för att inte få ett alltför spretigt och svårtolkat experiment har vi valt att ta bort det lägsta och högsta. Att vi valde just dessa

bottnar i att dessa två komprimeringsgrader var de minst representerade, vilket gör att vi kommer att använda oss av följande komprimeringar:

- 96 kbps
- 128 kbps
- 160 kbps
- 192 kbps
- 256 kbps

För att ta reda på huruvida vi skall använda oss av fast eller variabel bithastighet (för förklaring, se teoriavsnitt) har vi laddat ned ett antal av dessa filer. I stort sett alla av dessa använder sig av fast bithastighet vilket gör att vi kommer att göra det samma i vårt experiment.

Det program som vi har använt oss av för att skapa de olika mp3-filer som skall användas i experimentet är Steinbergs Wavelab. För att uppnå så bra kvalitet som möjligt har vi först spelat in ljudstycket på datorn i waveformat med 44 kHz samplingsfrekvens och med 16-bitars upplösning.

3.5.2. Val av ljudfiler

I experimentet kommer vi att använda oss av tre olika slags ljudfiler för att på det sättet kunna belysa om det finns några aspekter när det gäller komprimering av olika slags inspelningar. De tre kategorier vi kommer att använda oss av är:

- Tal. Valet av denna kategori grundar sig på att man inom e-learning och olika former av presentationer använder sig av inspelat tal. Här kommer vi att använda oss av Astrid Lindgrens välkända sagor som finns tillgängliga på CD. Valet av Astrid Lindgren grundar sig på ett bekvämlighetsval men också på att det i dessa sagor finns med en rad olika människoröster vilket medför att vi i testet har med både mans och kvinnoröster.
- Populärmusik. Många professionella ljudanvändare använder sig av Mp3-formatet för att kunna lagra och distribuera musik i ett smidigt format. De två låtar som skall användas är Anastacia, Paid my dues och Cher, Music is everything. Valet av dessa grundar sig på att dessa var de mest spelade på P3 under v.49/2001, som var veckan innan vi började med vår uppsats. Anledningen till att vi har valt ovanstående metod är att vi ville ha ett objektiva urval av musik som inte är influerat av vår egen musiksmak.
- Storskalig musik. För att kunna påvisa eventuella brister i komprimeringskvalité som inte kanske kommer fram i någon av de andra ovanstående kategorierna har vi valt att ta med klassiska musikstycken som har ett stort frekvensomfång och dynamik. Dessa musikstycken hämtades genom bekvämlighetsurval från en cd skiva som distribueras med tidningen High Fidelity. Återigen är anledningen till att vi valde dessa stycken är att vi inte ville att vår egen smak skulle påverka valet av musikstycke.

I och med att de tre ljudkategorierna spelas upp på två olika ljudsystem kommer vi att välja olika stycken för de olika systemen för att på det sättet minska chansen att respondenterna lär sig hur de olika styckena skall låta.

3.5.3. Val av uppspelningsordning

Uppspelningsordningen har vi fått fram genom ett slumpmässigt urval. Vi har låtit en tärnings utfall bestämma i vilken följd de olika komprimeringarna skall spelas upp. Vi gav de olika bithastigheterna värden på tärningen, det lägsta

bithastighetsvärdet, 96 kbps, blev 1 på tärningen, det näst lägsta 2 osv. Originalljudet fick värdet 6 på tärningen. Vi slog med tärningen tills vi fick fram en uppspelningslista med samtliga värden. Denna procedur upprepades för varje kategori och högtalaruppsättning. Ordningföljden inom ljudparen är även de framtagna genom slumpmässigt urval via tärning. I detta fall lät vi originalet få jämna värden och komprimeringen ojämna värden.

3.5.4. Val av uppspelningsutrustning

Experimentet kommer att använda sig av två olika ljudsystem för att på det sättet simulera olika ljudmiljöer. Det första systemet skall bestå av en standarddator med standardkomponenter för att på det sättet simulera en normal datoranvändares utrustning. Valet av detta system är en dator från högskolan Trollhättan/Uddevalla med medföljande komponenter. Anledningen till detta är dels av bekvämlighet i och med att vi har systemet lättillgängligt men också för detta system inte skiljer sig i allt för hög grad från andra system som olika leverantörer erbjuder.

Följande specifikationer gäller för den dator som används i experimentet.

- Pentium III 667 Mhz (Processor)
- 256 Mb RAM (Internminne)
- Quantum fireball P LM 18 Gb (Hårdisk)
- Soundblaster PCI-128 (Ljudkort)
- Trust Soundwave 40 (Högtalare)

Det andra systemet kommer att bestå av aktiva studiomonitorer (högtalare med inbyggd förstärkare avsedda för att användas i inspelningsstudios) och en vanlig Cd-spelare. Detta för att vi vill eliminera eventuell påverkan från en dators ljudkort, DA omvandlare och en vanlig Pc:högtalares begränsningar (Hacker, 2000). Valet av högtalare har fallit på ett par Genelec 1030 vilka kan ses som en form av industristandard när det gäller närfältsmonitorer (högtalare som placeras nära lyssningspositionen) för inspelningsstudios runt om i världen, vinnare av bl a TEC (Technical Excellence and Creativity) (Mediaspec, 2001). Dessa högtalare har inbyggda förstärkare vilket eliminerar behovet av någon extern vilket i sin tur gör att man eliminerar eventuella brister som kan uppstå i matchningen mellan högtalare och förstärkare. Dessa högtalare har följande specifikationer:

- Basskärning -3 dB: < 52 Hz
- Diskantskärning -3 dB: > 20 kHz
- Frekvensomfång 55 Hz - 18 kHz (± 2.5 dB)
- Maximalt ljudtryck @ 0.5m > 105dB SPL @ 1m > 111 dB SPL

Högtalaren har ett frekvensomfång på 55Hz-18 kHz med en maximal avvikelse på plus/minus 2,5 dB. Detta gör att högtalaren kommer att ha en helt neutral karaktär vilket i sin tur gör att avvikelser i uppspelningsmaterialet är lätta att höra.

Genom att vi kommer att ha två olika uppspelningssystem kommer vi kunna klargöra om det i systemens begränsningar finns intressanta aspekter att ta hänsyn till vid mp3-komprimering.

3.5.5. Experimentplats

Experimentet genomfördes i ett studierum i Uddevalla på Högskolan i Trollhättan Uddevalla. Rummet är ca 6*5 m, på den vägg dit ljudkällan varit riktad mot finns en Whitebordtavla i övrigt är det kala väggar. Ljudkällan har varit centralt placerad och respondenterna har varit placerade på rad ca 2,5 meter från ljudkällan. Då

rummet i vanliga fall fungerar som studierum eller grupprum består interiören av bänkar och stolar.

3.5.6. Förändringar efter testexperiment

Innan vi startade vårt experiment utförde vi ett test av det på en testgrupp. Detta gjordes för att vi ville ha synpunkter och feedback på vårt upplägg av experiment för att på detta sätt öka validiteten. Efter att vi hade utfört vår teststudie gjordes följande förändringar:

- Varje ljudstycke kortades ner från ca 30 till ca 15 sekunder i och med att vår testgrupp ansåg det svårt att komma ihåg hur föregående ljud lät vilket försvårade en rättvis bedömning av varje ljudpar.
- Vi minskade tystnaden mellan de två ljuden i ljudparet. Anledningen var densamma som ovannämnda.

3.6. Databehandling

Kvantitativ inriktning väljs oftast då man som vi söker kunskap skall kunna mätas eller beskriva och förklara. Därför kommer vi i vår analys och bearbetning välja ett statistiskt sätt att arbeta. (Patel & Tebelius 1987)

3.6.1. Svarsbearbetning

Efter att ha samlat in primärdata från experimentfasen kom nästa steg, att organisera och systematisera den kvantitativa datan. Vi valde att använda oss utav Excel. Syftet var att presentera data i en överskådlig och tolkningsbar form (Backman, 1998).

I nästa steg försökte vi tolka svaren och finna samband. Detta steg utförde vi var för sig för att sedan jämföra om vi hade grupperat svarsalternativen på samma sätt. De svar som vi hade skilda uppfattningar om diskuterade vi oss fram till en bra lösning. Dessa moment var tidskrävande och genomsyrades av ett iterativt arbete. Tanken med detta arbetsätt var att säkra reliabiliteten.

Vid inmatningen av rådata deltog båda i arbetet för att minska eventuella inmatningsfel i filen. Denna åtgärd vidtogs i syfte att minska bearbetningsfelen. Svaren sparades för att vi skulle ha möjlighet att gå tillbaka till ursprungsmaterialet. Båda arbetsmetoderna syftade till att erhålla högre reliabilitet.

3.7. Kvalitetsgranskning

För att stärka kvalitén i underökningen har vi i varje steg i studien haft validiteten och reliabiliteten i åtanke. För att belysa detta kan nämnas att med operationaliseringen som grund, utkristalliserade vi konkreta frågor med tydlig skärpa. I kombination med vår ansats dvs. att aktiviteterna i experimentet skall realiseras medför detta att respondenterna ger sakliga svar, vilket borgar för god kvalitet.

4. Teori

I det här stycket tar vi upp grunderna för vad ljud, digitalisering och komprimering är samt att vi förklarar olika begrepp som är relevanta till dessa.

4.1. Begreppsdefinitioner

I studiens teoridel diskuteras ett antal centrala begrepp som är viktiga för uppsatsen. Inom dessa begrepp förekommer det en hel del uttryck och ord som kommer att fungera som en teoretisk referensram.

4.1.1. Ljud

Begreppet ljud kan definieras på två olika sätt. Antingen som ljudvågor som rent fysikaliskt är mekaniska svängningar i ett elastiskt medium som kan vara i form av gas, vätska eller fast form. Eller som ljudförmimelse vilket innebär att ljudet förmedlas genom ett hörselorgan. (Arlinger & Lidén, 1985)

När vi hör något så är det ljudvågor som träffar örats trumhinna och sätter den i svängning och detta i sin tur stimulerar hörselnerverna. Ljudvågorna är i själva verket förändringar i lufttryck vid trumhinnan. (Nilsson, 1998b)

4.1.2. Frekvens

Antalet svängningar per sekund i ljudvågen kallas för frekvens och dess värde anger vilken ton ljudet har. Om frekvensen är hög (många svängningar) blir tonen hög (diskant) och omvänt om frekvensen är låg (få svängningar) blir tonen låg (bas). Frekvens mäts i Hertz. (Smelink, 2001)

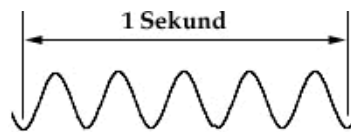


Bild 4.1 Frekvensomfång (Nilson, 1998).

4.1.3. Amplitud & Decibel

Amplituden beskriver graden av lufttrycksförändring vilket i sin tur visar sig genom vilken ljudstyrka som det aktuella ljudet har, se bild 4.2. Ljudtrycket mäts i decibel och där gränsen för hörbarhet är 0 dB och smärtgränsen för det mänskliga örat är 130 dB (Nilson, 1998). Den minsta ljudstyrkeförändringen som en människa kan uppfatta är en ökning eller minskning av ljudet med 3dB (Andersson, 2001).



Bild 4.2 Amplitud (Nilson, 1998)

4.1.4. Audio

Audio är ljud som återges av någon typ av media. Audio är ett akustiskt ljudomfång som är hörbart för människan (Whatis?.com, 2000).

4.1.5. Psykoakustik

Psykoakustik beskriver den mänskliga hörselns karakteristik. Detta grundar sig på att den mänskliga hörseln inte uppfattar frekvenser eller dynamik linjärt. Örat är känsligast mellan 2,5 kHz och 5 kHz, över och under dessa frekvenser sjunker

känsligheten Den viktigaste faktorn är dock maskeringseffekten vilket gör att varje ljud i en audiosignal täcker en viss del av närliggande frekvenser. Täckningsstorleken bestäms av hur starkt ljudet är. Styrkan på ljudet har också betydelse av hur starka ljud den kan täcka. Ett mycket starkt ljud kan täcka starka ljud i dess närhet men ju längre från ljudet man går desto svagare ljud kan den täcka. (Eriksson, 2001) Nedan visas ett exempel på detta:

Ex. Man har ett ljud 1000Hz med 90dB intensitet och ett annat ljud 1100Hz med 75 dB intensitet. 1100Hz ljudet kan inte höras för att 1000Hz ljudet täcker detta totalt. Om man däremot har ett ljud 2000Hz med intensiteten 75dB kan detta höras för att 1000Hz ljudet inte längre kan täcka det. (Eriksson, 2001)

Människans hörsel uppfattar inte ljudsignaler i nära varandra i tid linjärt. Efter ett starkt ljud (30dB-40dB skillnad) behöver människans hjärna en tid att uppfatta att ljud nivån har sjunkit. Människan behöver i medeltal 100ms tid att förstå att ljud nivån har sjunkit.(Eriksson, 2001)

4.2. Mänskligt hörande

En människas ljudomfång beror på flera faktorer bland annat ålder, ljudvolym och vilka skador örat har (Whatis?.com, 2000). Anledningen till att en människa uppfattar ljud beror vanligtvis på att en ljudvåg träffar trumhinnan och sätter denna i rörelse. Rörelsen överförs via hörselbenen (hammaren, städet och stigbygeln) till innerörat via basilarmembranet. Detta membran vibrerar och sätter vätskan i snäckan i rörelse och vilken upptas av de 23 500 hårceller som finns inuti snäckan. Hårcellerna omvandlar denna energi från mekanisk form till elektrisk (på samma sätt som i en mikrofon) och överför denna information till de nervtrådar som förbinder örat med hjärnan. (Hägerby, 1971)

En ljudupplevelse behöver nödvändigtvis inte vara en ljudvåg som träffar trumhinnan utan kan teoretiskt sett träffa vilken del av kroppen som helst och om den då har tillräckligt energiinnehåll sätta någon skelettdel i vibration som då överför detta till innerörat. (Hägerby, 1971)

Hörtröskeln för ljud, dvs. lägsta ljudtrycksnivå som kan uppfattas av en människa beror på vilken frekvens som ljudet har. Ljud som ligger inom talområdet kan uppfattas vid svagare ljudtryck än ljud som ligger i de undre registerna. (Arlinger & Lidén, 1985)

Förmågan att uppfatta ljud påverkas i många sammanhang av ljudsignalens längd. Detta innebär att ljud som har en varaktighet under 0,5 s måste vara starkare än de omgivande ljuden för att kunna höras. Det mänskliga örats känslighet avtar efter några minuters lyssning och medför att hörtröskeln försämras med ca 5-10 dB. (Arlinger & Lidén, 1985)

Maskering av ljud, dvs. att man har svårare att uppfatta ett ljud i närvaro av ett annat leder till att en signal som sträcker sig över ett större frekvensområde kan göra att det blir svårare att urskilja rena toner inom detsamma (Arlinger & Lidén, 1985). Maskeringen är mer effektiv i frekvensområdet som ligger över den maskerade tonen än i området under. Om intensitetsskillnaden är mindre än 20 dB mellan de båda tonerna blir maskeringen ofullständig. Denna maskering av ljud brukar oftast beskrivas med en psykoakustisk kurva. (Hägerby, 1971)

4.2.1. Frekvensområden

Den mänskliga hörseln kan uppfatta ljud mellan 20 Hz till 20 000 Hz. Det mänskliga talets totala frekvensområde ligger mellan 100 Hz och 8 000 Hz och där den viktigaste informationen sett ur ett talkommunikativt perspektiv återfinns mellan 300-3 000 Hz. (Arlinger & Lidén, 1985)

En vanlig uppdelning av det totala frekvensområdet brukar vara

- Bas, 20-200 Hz
 - Mellanregister 200-5 000 Hz
 - Diskant 5 000-20 000 Hz
- (Nyqvist & Fink, 1995)

4.3. Digitaliserat ljud

Vanligt ljud är av naturen analogt, medan digitaliserat ljud egentligen är ett samplat (inspelat) analogt ljud, som man använder olika metoder på för att kunna lagra och spara det på ett digitalt media. För att kunna lyssna på digitalt ljud krävs det att man återger ljudet via en analog källa (högtalare). Detta gör att man måste tillämpa olika metoder för att kunna omvandla ljudinformationen mellan analogt och digitalt. (Mülen, 1998) För att överhuvudtaget kunna beskriva hur det hela går till måste man först reda ut vissa begrepp

4.3.1. Samplat ljud

För att kunna sampla (spela in) ett ljud till datorn måste man först ha bestämt vilket frekvensområde som är aktuellt. Sedan måste man även ta hänsyn till vilken samplingsfrekvens, dvs hur många ljuprov som skall spelas in samt vilken upplösning inspelningen skall göras i, vilket förklaras nedan. (Mülen, 1998)

4.3.2. Samplingsfrekvens

När man skall sampla ett ljud måste man först bestämma med vilka tidsintervaller som amplituden av ljudet skall samplas i, (se bild 4.3). Man brukar tala om vilken samplingsfrekvens som används. En tumregel säger att samplingsfrekvensen bör vara minst dubbelt så stor som det analoga ljudets frekvens. För att uppnå Cd-kvalitet på ljudet måste samplingsfrekvensen vara 44 kHz. (Nilsson, 1998)

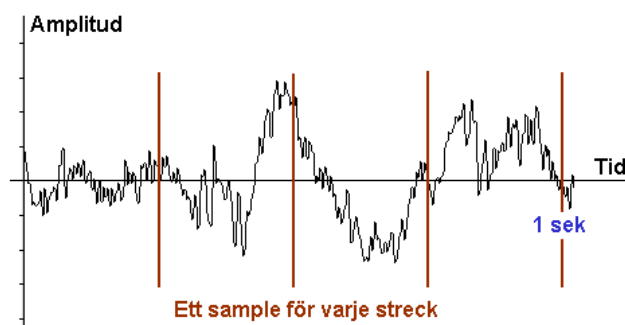


Bild 4.3 samplingsfrekvens 4 ggr per sekund (Nilsson, 1998).

4.3.3. Bitdjup & upplösning

När tidsintervallet för samplingen är bestämt måste man också ta hänsyn till hur noggrann avläsningen av amplituden inom intervallet skall vara. Detta bestäms utifrån hur många minnesbitar som man sparar informationen i. Ju fler minnesbitar som används desto fler nivåer av amplituden kan sparas vilket gör att

nyansrikedomen i ljudet ökar (Mülen, 1998). I bild 4.4 visas en sampling med 8-bitar vilket gör att man kan beskriva 256 styrkenivåer i ljudet (Nilsson, 1998).



Bild 4.4, 8-bitars sampling (Nilsson, 1998).

När man har tagit allt detta i beaktning kan man börja att spela in ljudet till datorn. För att kunna beräkna hur stor plats det inspelade ljudet kommer att ta upp kan man använda sig av följande formel:

Samplingsfrekvens x ljudets längd i sekunder x (bitdjup/8) x 2(för stereo)

Om man vill spela in en minut ljud i Cd-kvalitet blir formeln följande:

44100 x 60 x 16/8 x 2 ger ett ljud som tar upp = 10 MB (10 584 000 byte)

(Nilsson, 1998)

4.4. Komprimering

Komprimering av data görs för att man vill minska överföringstid och spara plats på lagringsmedia. Själva komprimeringen kan göras på två olika sätt, förlustfri och förlustgivande. Vid komprimering av ljud medför detta oftast en kompromiss mellan kvalitet och överföringshastighet. (Nilsson, 2000)

För att kunna utföra en komprimering av ljud behövs det två saker. Dels en encoder som utför själva ”packningen” av ljuddata. Encodern är i grunden en matematisk algoritm som analyserar och beräknar inkommande data för att sedan kunna representera materialet på ett mer kompakt sätt. Denna process är den som kräver de största dataresurserna när det gäller komprimering. För att kunna spela upp det komprimerade ljudet behövs det en decoder som är en process genom vilket datainformationen packas upp som skapats av encodern. (Andersson & Sönnebo, 1998)

4.4.1. MPEG

MPEG står för Moving Picture Expert Group och refererar till en metod av data kompression som är optimerad för video och ljudformat. Filerna måste alltid dekomprimeras före användning. MPEG-2 är en efterföljare till MPEG-1. Oftast används MPEG-1 för CD-ROM eller VideoCD och MPEG-2 används för utsändning eller DVD. En uppmärksam skillnad mellan 1:an och 2:an är att MPEG-2 har implementerat varierande samplingsfrekvens. (Jansson, 2000)

Mpeg-komprimering är en effektiv metod som kan ta bort mycket information men ändå behålla stor del av ljudkvaliteten. Denna metod finns i flera olika utförande men här tas endast Mpeg.1 layer 3, dvs. mp3 upp. (Nilsson, 1998)

4.5. Mp3

Okomprimerat ljud likt det på Cd-skivor innehåller mer data än vad våra hjärnor kan uppfatta. Exempelvis så är det så att om man spelar upp två ljud med samma frekvens men med olika ljudnivå så uppfattar vår hjärna enbart det högre ljudet. Vidare har vi svårt att uppfatta två toner om de är väldigt lika varandra. Örat är också känsligare för vissa toner än för andra. När man skapar mp3 analyserar "encoding tools" den inkommande signalen och bryter ner den till matematiska mönster som senare jämförs med psykoakustiska modeller som finns i verktyget. Verktyget tar bort de data som inte matchar de givna modellerna för mänskligt hörande. Komprimeringen är förstörande, vilket innebär att viss data tas bort i komprimeringen. Mp3-filer tar därför alltid bort viss information från originalljudet oavsett kvalitet. (Nielsen, 2001)

Att mp3 tar bort data i komprimeringen gör i vissa fall inte så mycket, då människan ändå inte kan uppfatta dessa dataförluster. Även om detta är väletablerade fakta om hörande är hjärnan som styr det hela lite mer komplicerad än så. Med detta menas att hjärnan reducerar intryck och enbart tar åt sig det mest relevanta. Det finns de som uppskattar att vi egentligen bara tar emot en miljarddel av de data som är mottagbar för våra fem sinnen vid vilken tidpunkt som helst. Därför kan man säga att en av hjärnans viktigaste funktioner är att skilja mellan relevant och icke-relevant data. (Hacker, 2000)

Därför är den grundläggande tanken inom komprimering att ta bort de data vi ändå inte uppfattar. Detta låter självklart men sanningen är att bra "inspelningar" innehåller mer ljud data än vi kan ta emot därför att inspelningsutrustningen har större ljudomfång än det mänskliga örat. Mp3 kallas på engelska för ett "perceptual codec", vilket innebär att det komprimerar enligt matematiska uträkningar av det mänskliga hörandet. Mp3 är på intet sätt det ultimata ljudalternativet. Men verkligheten är att de flesta användare inte har det tränade örat att upptäcka skillnaderna från originalet eller för den delen innehar den bästa utrustningen. Vi har kanske inte hur mycket utrymme som helst på våra hårddiskar eller de bästa ljudkortet och högtalarna till vår Pc:anläggning. Faktum är att kvalitén på dessa komponenter ligger långt efter kvalitén på stereoanläggningen vi har där hemma gör att vi i många fall ändå inte uppfattar skillnaden. (Hacker, 2000)

4.5.1. Bithastighet

Vid Mp3-komprimering finns det två saker som påverkar processen. Dels är det den matematiska jämförelsen med den psykoakustiska modellen som finns inbyggd i encodern och dels den bithastighet som man väljer innan själva komprimeringen startar. Bithastigheten refererar till de antal bitar per sekund som skall användas för att spara den slutgiltiga produkten. Ju högre bithastighet som används desto större upplösningen av ljudet. För att beskriva bithastighetens inverkan på ljudkvaliteten kan man jämföra det med gamla stumfilmer som ofta vara "hoppiga". Detta berodde på att man använde sig av mindre antal bildrutor per sekund vid inspelningen vilket i sin tur medförde att mindre information distribuerades över en given tid. (Hacker, 2000)

Vid komprimering finns det två olika sorters bithastigheter som man kan använda sig av. CBR (Constant Bit Rate) betyder att man får den upplösning på ljudet som man har angivit. Nackdelen med detta är att musik inte följer en konstant upplösning utan kan variera såsom partier med mycket musik som bryts av med partier med bara sång. På grund av detta har man utvecklat encoders som stödjer

VBR (Variable Bit Rate) vilket innebär att encodern varierar bithastigheten beroende på dynamiken i signalen. Denna variant har också en del nackdelar, främst när det gäller kompatibiliteten med gamla decoders och att tidsfel kan uppstå vid uppspelning av filen. (Hacker, 2000)

Bithastighet skall inte förväxlas med samplingsfrekvens (samplerate). Bithastighet beskriver storleken på data som sparas för varje sekund medan samplingsfrekvens anger den frekvens som inspelningen sparas med. (Hacker, 2000)

4.5.2. Filstorlek

Vitsen med att använda sig av komprimering är att minska lagringsutrymmet. En Mp3-fil kan variera i storlek beroende på vilken bithastighet som används för komprimeringen. I nedanstående tabell visas de förhållande som finns mellan olika bithastigheters och filstorlek. (Filen som används är 3:52 minuter lång)

Bithastighet (kbps)	Filstorlek (i megabyte)	Kompressionsgrad
Original fil	40,08	1:1 (ingen)
16	0,46	87:1
32	0,92	44:1
64	1,83	22:1
96	2,74	15:1
128	3,64	11:1
160	4,55	9:1
192	5,46	7:1
256	7,28	6:1
320	9,10	4:1

Tabell 4:1 Förhållande mellan komprimeringsgrad och filstorlek. (Hacker, 2000)

4.5.3. Ljudkvalitet

I och med att Mp3 är en förstörande komprimering står ljudkvaliteten i proportion till hur mycket information som tas bort. 128 kbps är den lägsta nivån man skall använda sig av när man komprimerar musik. Skall musikfilerna användas seriöst är det att fördra 160 eller 192 kbps i bithastighet. (Hacker, 2000)

Enligt Fraunhofer Institut uppnår man följande kvalitet på ljudet vid olika bithastigheters enligt nedanstående tabell.

Ljudkvalitet	Frekvensomfång	Bithastighet
Telefonljud	2,5 kHz	8 kbps
Bättre än kortvågsradio	4,5 kHz	16 kbps
Bättre än AM radio	7,5 kHz	32 kbps
Ungefär som FM radio	11 kHz	64 kbps
Nära CD ljud	15 kHz	96 kbps
CD-kvalitet	>15 kHz	128 kbps

Tabell 4:2 Ljudkvalitet enligt Fraunhofer (Fraunhofer, 2000)

5. Resultat

I resultatdelen redovisar vi våra svar i form av enkla tabeller. Varje ruta innehåller ett ljudpar och i rubriken framgår det vilken komprimering som behandlas. För enkelhetens skull har vi valt att redovisa komprimeringarna i en stegrande storleksordning med det lägsta bithastighetstalet först och parett med de båda originalljuden sist. I rubriken framgår det också i vilken ordning komprimeringen och originalljudet spelades upp.

I varje ruta redovisas våra två respondentgrupper uppdelade var för sig, där vi gett urvalsgruppen med personer som kommer i kontakt med ljud i undervisningssyfte i sin skol- eller arbetsmiljö för *Studenter* och urvalsgruppen som i sitt arbete eller på en stor del av sin fritid kommer i kontakt med musik samlingsnamnet *Ljudproffs*. Dessa samlingsnamn har ingen egentlig betydelse utan är tänkta att underlätta för oss själva och läsarna av denna uppsats. Slutligen redovisas även en sammanslagning av de två respondentgrupper under namnet *Alla*. Vidare behöver man förstå följande begrepp för att kunna tolka våra tabell på ett rättvisande sätt:

- *Lika*, respondenten uppfattade inte någon skillnad mellan de två uppspelade ljuden
- *Org*, respondenten tyckte att originalljudet lät bättre än komprimeringen.
- *Komp*, respondenten uppfattade det komprimerade ljudet som det bättre.

5.1. Åldersgrupper

På vår fråga om vilken åldersgrupp man tillhörde fick vi följande svar:

- 15-25 år, 12 st
- 26-35 år, 16 st
- 36-45 år, 2 st
- 45<, 0 st

5.2. Tal på PC högtalare

Här nedan redovisas undersökningens resultat på tal på PC högtalarna i tabellform. Ljudfilen som spelades var en del ur Astrid Lindgrens *Emil i Lönneberga*. Resultatet redovisas även i diagramform i Bilaga 1.

PC-Tal, Original-96 kbps				PC-Tal, 128 kbps-Original				PC-Tal, Original-160 kbps			
	<i>Lika</i>	<i>Org</i>	<i>Komp</i>		<i>Lika</i>	<i>Org</i>	<i>Komp</i>		<i>Lika</i>	<i>Org</i>	<i>Komp</i>
<i>Studenter</i>	6	3	6	<i>Studenter</i>	6	2	7	<i>Studenter</i>	12	2	1
<i>Ljudproffs</i>	8	2	5	<i>Ljudproffs</i>	9	4	2	<i>Ljudproffs</i>	10	1	4
<i>Alla</i>	14	5	11	<i>Alla</i>	15	6	9	<i>Alla</i>	22	3	5
PC-Tal, Original-192 kbps				PC-Tal, 256 kbps-Original				PC-Tal, Original-Original			
	<i>Lika</i>	<i>Org</i>	<i>Komp</i>		<i>Lika</i>	<i>Org</i>	<i>Komp</i>		<i>Lika</i>	<i>Första</i>	<i>Andra</i>
<i>Studenter</i>	13	1	1	<i>Studenter</i>	8	3	4	<i>Studenter</i>	6	3	6
<i>Ljudproffs</i>	9	3	3	<i>Ljudproffs</i>	7	2	6	<i>Ljudproffs</i>	3	4	8
<i>Alla</i>	22	4	4	<i>Alla</i>	15	5	10	<i>Alla</i>	9	7	14

Tabell 5.1 Tal på PC högtalare

5.3. Populärmusik på PC högtalare

Här nedan redovisas resultatet av populärmusiken på PC högtalarna. Musiken som spelades upp för respondenterna var Chers *Music is everything*. Resultatet redovisas även i diagramform i Bilaga 2.

Pcmusik 96kbps-Orig				Pcmusik, Orig-128 kbps				Pcmusik, orig-160kbps			
	<i>Lika</i>	<i>Org</i>	<i>Komp</i>		<i>Lika</i>	<i>Org</i>	<i>Komp</i>		<i>Lika</i>	<i>Org</i>	<i>Komp</i>
<i>Studenter</i>	6	6	3	<i>Studenter</i>	7	6	2	<i>Studenter</i>	7	6	2
<i>Ljudproffs</i>	4	10	1	<i>Ljudproffs</i>	8	4	3	<i>Ljudproffs</i>	8	5	2
<i>Alla</i>	10	16	4	<i>Alla</i>	15	10	5	<i>Alla</i>	15	11	4
Pcmusik, Orig-192kbps				Pcmusik, 256kbps-Orig				Pcmusik, Orig-Orig			
	<i>Lika</i>	<i>Org</i>	<i>Komp</i>		<i>Lika</i>	<i>Org</i>	<i>Komp</i>		<i>Lika</i>	<i>Första</i>	<i>Andra</i>
<i>Studenter</i>	7	7	1	<i>Studenter</i>	7	5	3	<i>Studenter</i>	10	2	3
<i>Ljudproffs</i>	5	4	6	<i>Ljudproffs</i>	8	5	2	<i>Ljudproffs</i>	7	4	4
<i>Alla</i>	12	11	7	<i>Alla</i>	15	10	5	<i>Alla</i>	17	6	7

Tabell 5.2 Populärmusik på PC högtalare

5.4. Klassisk musik på PC högtalare

Nedanför redovisas resultaten av den klassiska musiken på PC högtalarna. Ljudfilen var ett stycke ur Mendelsons symfoni nummer 2 *Lobgesang*. Resultatet redovisas även i diagramform i Bilaga 3.

PC-klas Orig-96kbps				PC-klas 128kbps-Orig				PC-klas, 160kbps-org			
	Lika	Org	Komp		Lika	Org	Komp		Lika	Org	Komp
Studenter	9	4	2	Studenter	10	3	2	Studenter	7	3	5
Ljudproffs	11	1	3	Ljudproffs	6	2	7	Ljudproffs	5	6	4
Alla	20	5	5	Alla	16	5	9	Alla	12	9	9
PC-klas Orig-192kbps				PC-klas, 256kbps-org				PC-klas Orig-Orig			
	Lika	Org	Komp		Lika	Org	Komp		Lika	Första	Andra
Studenter	8	3	4	Studenter	10	2	3	Studenter	10	4	1
Ljudproffs	7	4	4	Ljudproffs	8	4	3	Ljudproffs	5	6	4
Alla	15	7	8	Alla	18	6	6	Alla	15	10	5

Tabell 5.3 Klassisk musik på PC högtalare

5.5. Tal på studiomonitorer

I tabell 5.4 redovisas tal på studiomonitorerna i tabellform. Ljudfilen var en del ur *Emil i Lönneberga* av Astrid Lindgren. Resultatet redovisas även i diagramform i Bilaga 4.

Studiotal org-96kbps				Studiotal 128kbps-org				Studiotal org-160 kbps			
	Lika	Org	Komp		Lika	Org	Komp		Lika	Org	Komp
Studenter	9	4	2	Studenter	2	8	5	Studenter	7	4	4
Ljudproffs	7	1	7	Ljudproffs	9	3	3	Ljudproffs	5	4	6
Alla	16	5	9	Alla	11	11	8	Alla	12	8	10
Studiotal 192 kbps-org				Studiotal 256kbps-Orig				Studiotal Org-Orig			
	Lika	Org	Komp		Lika	Org	Komp		Lika	Första	Andra
Studenter	10	1	4	Studenter	10	1	4	Studenter	7	4	4
Ljudproffs	5	2	8	Ljudproffs	10	3	2	Ljudproffs	6	6	3
Alla	15	3	12	Alla	20	4	6	Alla	13	10	7

Tabell 5.4 Tal på studiomonitorer

5.6. Populärmusik på studiomonitorer

I tabell 5.5 redovisas resultaten av populärmusiken på studiomonitorerna. Musiken som spelades upp för respondenterna var Anastacias *Paid my dues*. Resultatet redovisas även i diagramform i Bilaga 5.

Studiosus org-96kbps				Studiosus 128kbps-org				Studiosus 160kbps-org			
	Lika	Org	Komp		Lika	Org	Komp		Lika	Org	Komp
Studenter	3	9	3	Studenter	9	5	1	Studenter	6	8	1
Ljudproffs	2	11	2	Ljudproffs	7	6	2	Ljudproffs	5	6	4
Alla	5	20	5	Alla	16	11	3	Alla	11	14	5
Studiosus org-192kbps				Studiosus 256kbps-org				Studiosus org-org			
	Lika	Org	Komp		Lika	Org	Komp		Lika	Första	Andra
Studenter	10	1	4	Studenter	7	6	2	Studenter	7	3	5
Ljudproffs	6	4	5	Ljudproffs	6	5	4	Ljudproffs	5	8	2
Alla	16	5	9	Alla	13	11	6	Alla	12	11	7

Tabell 5.5 Populärmusik på studiomonitorer

5.7. Klassisk musik på studiomonitorer

I denna tabell redovisas resultaten av den klassiska musiken på studiomonitorerna. Musikfilen som spelades upp för respondenterna var Romberg *opus 41* symfoni nummer 1. Resultatet redovisas även i diagramform i Bilaga 6.

Studioklas 96kbps-org				Studioklas org-128kbps				Studioklas org-160kbps			
	Lika	Org	Komp		Lika	Org	Komp		Lika	Org	Komp
Studenter	6	9	0	Studenter	8	3	4	Studenter	5	4	6
Ljudproffs	1	13	1	Ljudproffs	6	6	3	Ljudproffs	6	3	6
Alla	7	22	1	Alla	14	9	7	Alla	11	7	12
Studioklas org-192kbps				Studioklas 256kbps-org				Studioklas org-org			
	Lika	Org	Komp		Lika	Org	Komp		Lika	Första	Andra
Studenter	6	6	3	Studenter	11	3	1	Studenter	11	1	3
Ljudproffs	7	1	7	Ljudproffs	6	4	5	Ljudproffs	10	2	3
Alla	13	7	10	Alla	17	7	6	Alla	21	3	6

Tabell 5.6 Klassisk musik på studiomonitorer

6. Resultatdiskussion

I detta kapitel kommer vi att analysera de resultat som vi har kommit fram till i vårt experiment samt analysera vår metod. För att underlätta detta har vi tagit med ett antal diagram som åskådliggör resultatet. Alla diagram som används i vår diskussion har följande struktur:

- Värdeaxeln (Y) visar respondentsvar i antal eller procent.
- Kategoriaxeln (X) visar vilken bithastighet det är frågan om och samt vilket uppspelningsmedium som avses.
- Kurvorna respektive staplarna visar vad de olika respondenterna ansåg lät bäst.

6.1. Metoddiskussion

I vårt experiment finns vissa svagheter som gör att vårt resultat inte kan tolkas som hundra procentigt rättvisande svar på vår frågeställning. Vi har valt att tolka dessa utifrån tre olika perspektiv, respondenter, utrustning och plats, samt utförande.

6.1.1. Respondenter

I vår uppsats ville vi ha två urvalsgrupper där den ena var studenter som kommer i kontakt med ljud. Detta gjorde att det stora flertalet från denna grupp genom snöbollsmetoden kom från samma utbildning som oss. Att studenter på digital informationsproduktion passar i vårt urval är helt klart. Den svaghet som detta kan innebära är att i och med att respondenterna känner oss sen tidigare kan detta påverka deras svar.

Det har varit svårt att urskilja klara mönster i våra experimentsvar, det hade varit önskvärt med fler antal respondenter i varje urvalsgrupp. Genom att ha fler respondenter tror vi att det hade varit lättare att få fram och urskilja tydliga skillnader i experimentet. Med så få respondenter i varje urvalsgrupp går det inte att dra några som helst generella slutsatser utan de svar vi får fram kan endast ses som en fingervisning.

Förutom ålder har vi inte alls berört våra respondenters eventuella hörselnedsättning, sjukdom eller liknande som kan påverka deras uppfattning om hur de upplever de olika experimentljuden. Om vi hade använt oss av dessa

uppgifter och gjort jämförelser med det resultat som vi fått i experimentet hade vi fått alltför många variabler att ta hänsyn till för att kunna få ett tolkningsbart resultat, i alla fall för vår ansats.

6.1.2. Experimentutrustning och plats

Vårt val av utrustning grundar sig i viss mån på bekvämlighetsurval genom att det var det som vi hade tillgång till vid experimenttillfället. Det hade varit önskvärt att ha en tidigare studie av den utrustning som används i experimentet och framförallt när det gäller pc:högtalarna.

Experimentet utfördes i ett studierum på högskolan. I och med att vi inte har haft någon möjlighet att analysera hur rummet påverkar de uppspelade ljuden. Detta gör att vi inte kan säga hur stor inverkan rummet har haft på vårt resultat.

6.1.3. Utförande av experiment

Vid varje tillfälle var det mellan en och fyra respondenter. Att vi tillät så pass många vid experimenttillfället kan ha påverkat det resultat som vi har kommit fram till. Detta på grund av att respondenterna dels kan påverka varandra och dels att uppfattningen av ljudet kan skilja sig beroende på hur man sitter i förhållande till uppspelningsskällan.

6.2. Åldersgrupper

På vår enkät som användes under experimentet svarade respondenterna på vilken åldersgrupp de tillhörde. I och med att vi endast vid två fall hade respondenter utanför de två yngsta åldersgrupperna, som i sig hade en relativt jämn fördelning, har vi valt att inte beröra denna faktor något mer i vår uppsats.

6.3. Respondenter

Den enda tydliga skillnaden mellan våra urvalsgrupper är att i de flesta fall upplever ljudproffsen att dom hör skillnad mellan ljudparen i högre grad än studenterna. Denna skillnad gäller oavsett om man hör ”rätt”, dvs. att man hör vilket som är original eller inte. Detta mönster går även igen när det gäller original mot original, där ljudproffsen upplever i högre utsträckning än studenterna att dom hör skillnad. Man kan också se att ljudproffsen i större grad anser att det komprimerade ljudet låter bättre än originalet, detta gäller endast vid de högre komprimeringarna, dvs. 192 kbps och uppåt. I övrigt så har vi inte hittat några relevanta avvikelser som kan bilda ett genomgående mönster i svaren mellan de två respondentgrupperna. Vid något enstaka fall har vi påträffat skillnader men i och med att dessa inte är genomgående i vårt experiment har vi valt att inte lägga någon större vikt vid detta.

6.4. Tal på Pc:högtalare

I undersökningen över tal på PC högtalarna visar det sig att respondenterna i majoritet tycker att komprimering och original låtit lika och får också flest svar vid samtliga ljudpar, se diagram 6.1. Vid ljudparet 128 kbps komprimering mot original uppnådde originalet sitt högsta värde med 6 svar, alltså 20% av svaren, se diagram 6.1. I vår undersökning kunde alltså 20% eller mindre av respondenterna skilja på originalet gentemot komprimering. Anmärkningsvärt kan tyckas att den lägsta siffran av respondenter som tycker att ljudparen låter lika återfinns i jämförelsen mellan de båda originalen, se bilaga 1.

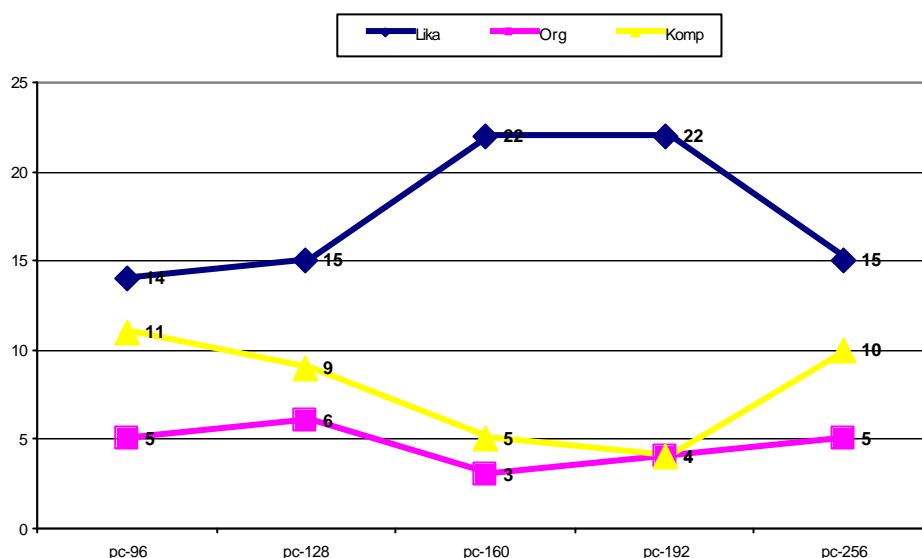


Diagram 6.1 Tal på pc:högtalare

6.5. Populärmusik på Pc:högtalare

I undersökningen om populärmusik på PC högtalarna visade det sig att komprimeringen genomgående fick lägst antal svar när man slog ihop de bägge gruppernas svar. Svaret lika var återigen det vanligaste alternativet, även om en majoritet upplevde originalet som bäst i ljudparet 96 kbps mot original, se diagram 6.2.

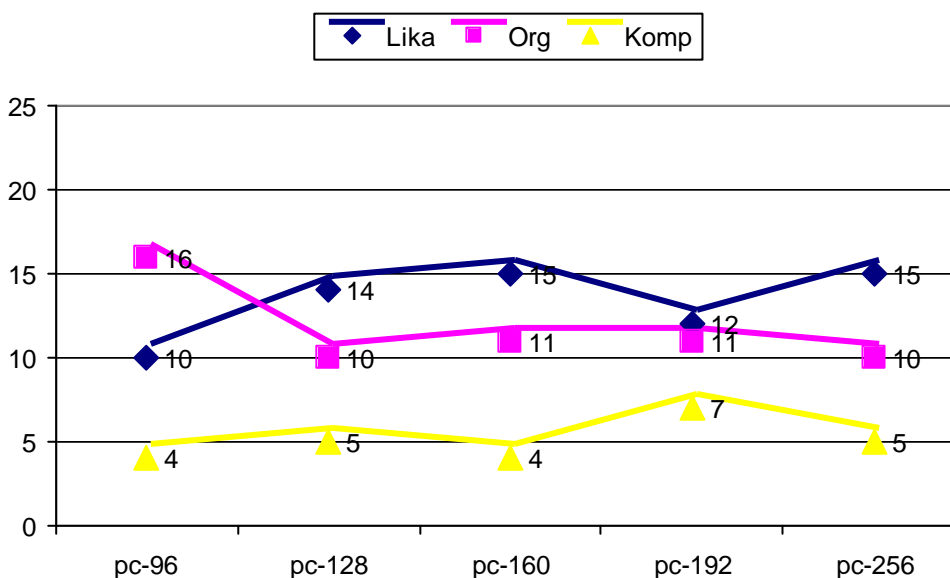


Diagram 6.2 Musik på pc:högtalare

6.6. Klassisk musik på Pc:högtalare

Angående klassisk musik på PC högtalarna i vår undersökning får alternativet att komprimering och original låter likadant genomgående flest antal svar. Originalen och komprimeringen följer varandra åt genom hela kurvan utan några större differenser dem emellan. Intressant fakta är att svaren med lika når sin topp i ljudparet original mot 96 kbps, samtidigt som originalet når sin lägsta svarsfrekvens vid samma ljudpar, se diagram 6.3.

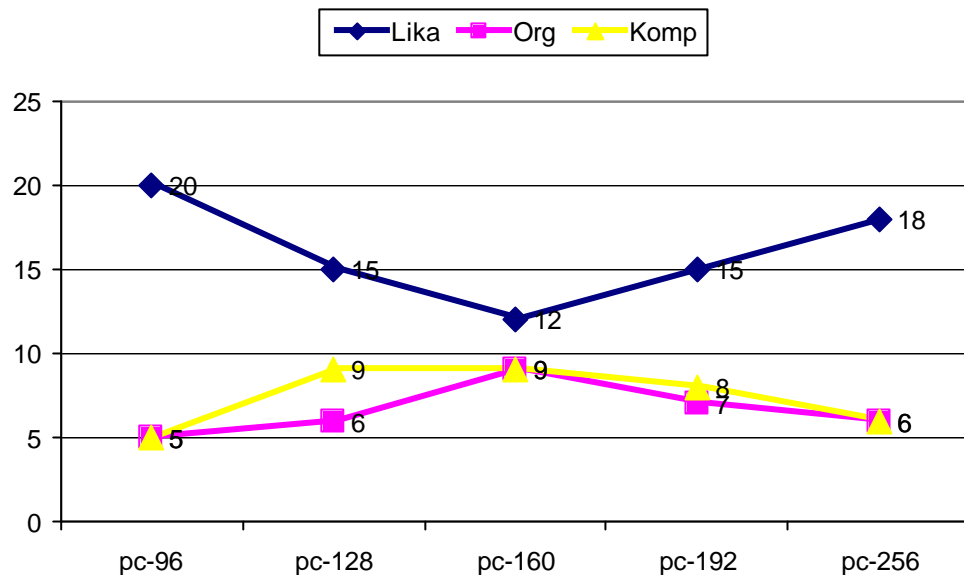


Diagram 6.3 Klassisk musik på pc: högtalare

6.7. Tal på studiomonitorer

Vid tal på studiomonitorerna visar det sig återigen att komprimering och original upplevs lika då det genomgående är det vanligaste svarsalternativet. Vid ett tillfälle sticker originalet ut, vid ljudparet 128 kbps och original får det lika många svar som alternativet lika och passerar också vid det tillfället komprimeringsalternativet, se diagram 6.4, som annars har något högre svarsfrekvens än originalet. Vid komprimeringsgrader över 160 kbps är det endast ett fåtal i vår undersökning som upplever originalet bättre än komprimering.

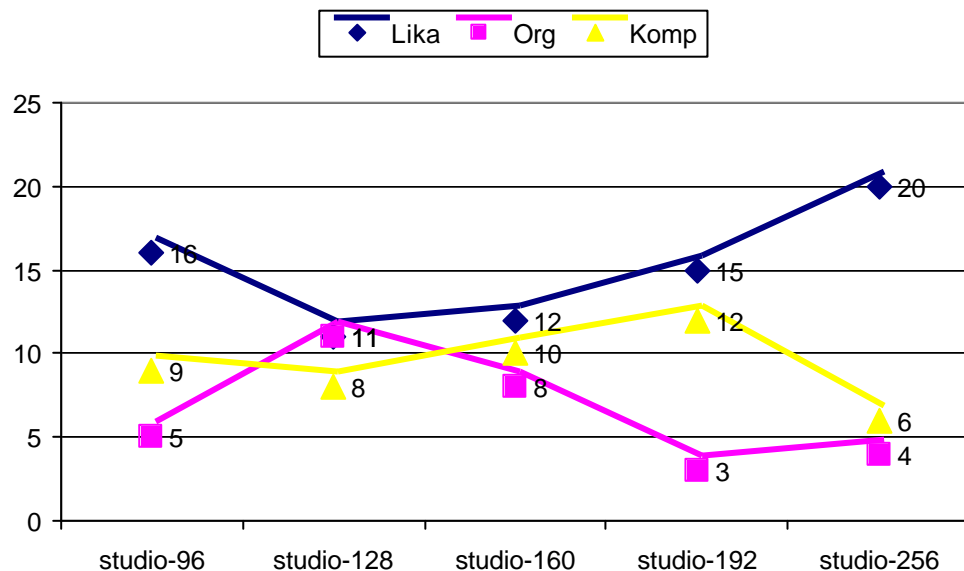


Diagram 6.4 Tal på studiomonitor

6.8. Populärmusik på studiomonitorer

Vårt experiment visar att det stora flertalet respondenter upplever att originalet låter bättre vid jämförelse med 96 kbps komprimering, se diagram 6.5. Vid komprimeringsgraderna 128 kbps och 192 kbps är det mer än hälften som upplever ljudet likvärdigt originalet.

Vid 192 kbps komprimering var det dessutom nio respondenter som upplevde komprimeringen som det bättre ljudet. Vilket resulterar i att det endast är fem respondenter som upplever att originalet låter bäst, se diagram 6.5.

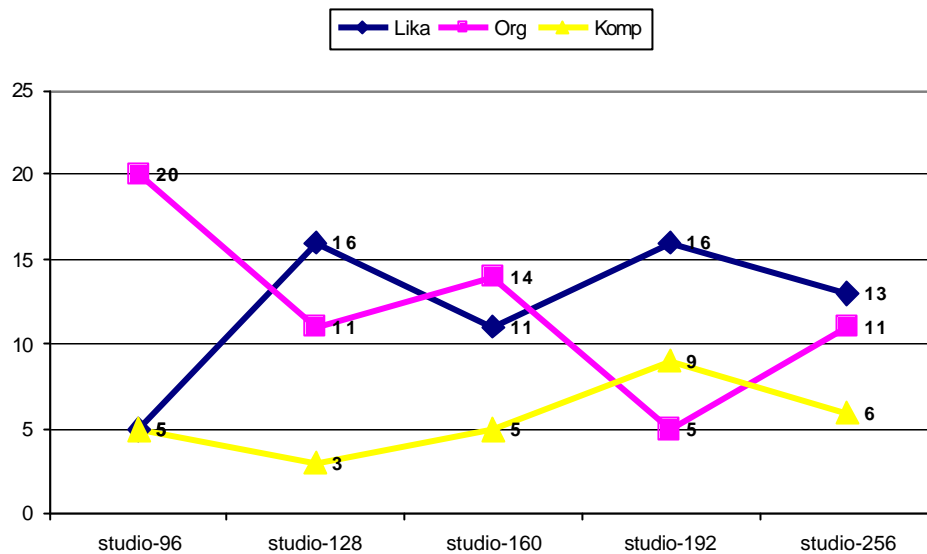


Diagram 6.5 Populärmusik på studiomonitorer

6.9. Klassisk musik på studiomonitorer

En klar majoritet av respondenterna i vårt experiment upplevde originalet bättre än 96 kbps. Tendensen för klassisk musik tycks vara att från 160 kbps och högre så är det en övervägande del som upplever det komprimerade ljudet som lika bra eller bättre än originalet, se diagram 6.6.

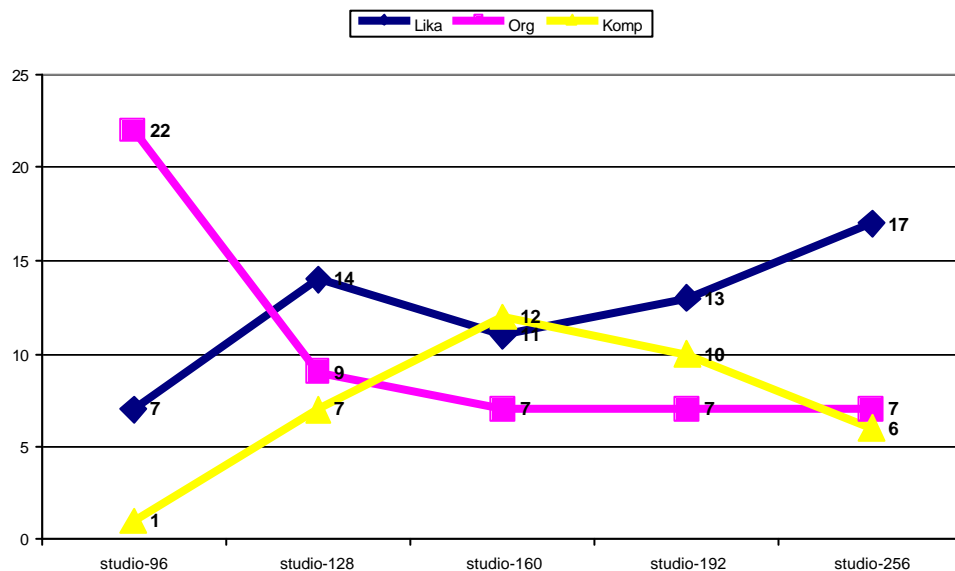


Diagram 6.6 Klassisk musik på studiomonitorer

6.10. Komprimerat tal

Vid 6 av 8 mättillfällen är det endast ett fåtal som upplever originalljudet som bättre än komprimeringen. Vid de andra två tillfällena är det endast 1/3 av respondenterna som upplever originalljudet som det bättre. Detta gäller för komprimeringsgraderna 128 & 160 kbps uppspelat på studiomonitorerna, se diagram 6.7.

I och med att det mänskliga talet ligger mellan 100 Hz och 8000 Hz och de flesta respondenter inte tycker att originalet låter bättre kan vi dra slutsatsen att komprimering av detta frekvensområde inte påverkas vid valet av bithastighet om man har tänkt sig att spela upp ljudet på pc:högtalare. För att säkerställa bästa ljudkvalitet om uppspelning skall ske på ett mer avancerat ljudsystem kan dock en bithastighet över 160 kbps vara att föredra. Detta medför att det är möjligt att distribuera tal för uppspelning på dator med en kompressionsgrad av 15:1. När det gäller undantagen vid tal på studiomonitor kan man fundera på om encodern har vissa brister när det gäller att komprimera tal vid dessa bithastigheter.

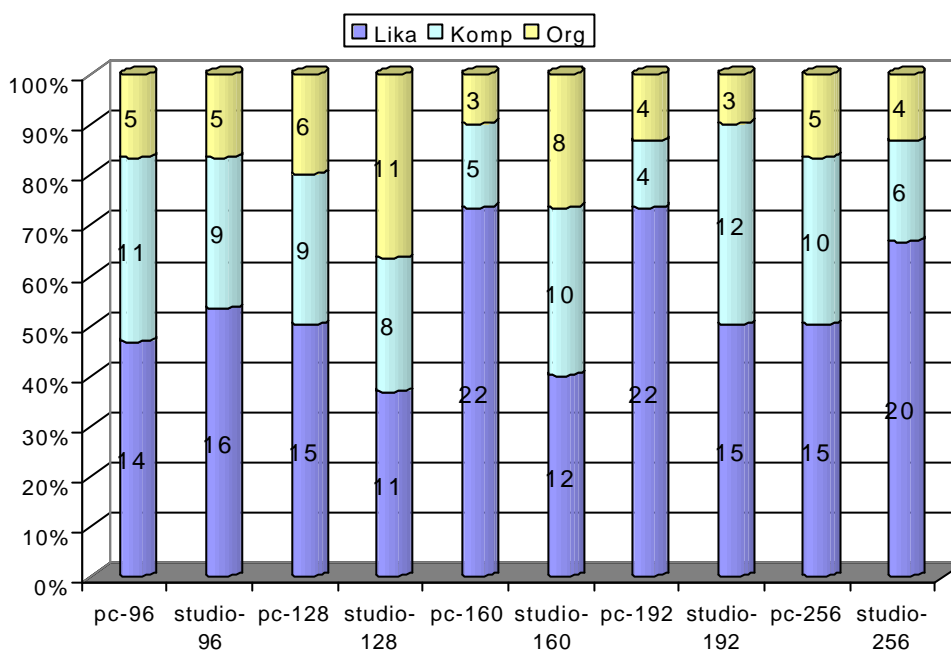


Diagram 6.7 Komprimerat tal på både pc och studio högtalare

6.11. Komprimerad populärmusik

Experimentet visar att flertalet av respondenterna tyckte att originalet lät bättre än komprimeringen 96 kbps. Detta gäller oavsett högtalarsystem. I sju av åtta fall vid de övriga komprimeringarna oavsett system är det ca 1/3 som upplever originalet som det bättre alternativet. Vid 192 kbps komprimering uppspelat på studiomonitorer är det endast 5 av 30 som upplever att originalet låter bättre, se diagram 6.8. detta stärker Hackers åsikt om att man minst komprimera i 160 kbps eller 192 kbps för att kunna använda musiken seriöst. Däremot kan Fraunhofers uppfattning att 128 kbps räcker för att uppnå Cd-kvalitet ifrågasättas.

Resultatet visar att bristerna som uppkommer vid komprimering av musik är lättare att avslöja på studiomonitorerna samt att det tydligare avspeglar sig att själva encoderns kapacitet skiljer sig mellan olika komprimeringsgrader. Detta är däremot inte så framträdande på pc:högtalarna vilket kanske kan hänföras till deras begränsade frekvensomfång samt de eventuella brister som datorns ljudkort kan ha.

I och med att det mänskliga örat kan höra mellan 20-20000 Hz enligt Arlinger & Lidén betyder detta att de brister som uppkommer i komprimering framförallt återspeglar sig i diskantområdet. Detta kan man dra som slutsats i och med att när det gällde tal som ligger mellan 100-8000 Hz i mycket mindre grad upplevs som sämre än musik som återfinns i hela det mänskliga hörselomfånget. Det här påståendet bygger framförallt på när vi spelade på våra studiomonitorer som har ett

frekvensomfång på 55Hz-18 kHz. Vi har inte några specifikationer på Pc:högtalarna som användes i experimentet men i och med att skillnaderna inte var lika tydliga på dessa kan man ana att de inte har ett lika stort frekvensomfång.

Jämfört med den klassiska musiken har de populärmusikaliska låtarna i vår undersökning en ljudbild som innehåller fler instrument. I vårt fall är det ju som så att våra respondenter hör mycket mer skillnad på populärmusiken än den klassiska. Detta kan knytas till Erikssons genomgång av psykoakustik samt Hackers och Nielsens förklaring av mp3-komprimeringens matematiska koppling till den psykoakustiska modellen som ligger till grund för vad som tas bort vid komprimering. I och med att man tar bort ljud som ligger nära varandra i frekvens och ljudtryck kan det vara så att man lättare hör en försämring av den komprimerade populärmusiken på grund av den komplexa audiosignalen med många instrument inom samma frekvensomfång.

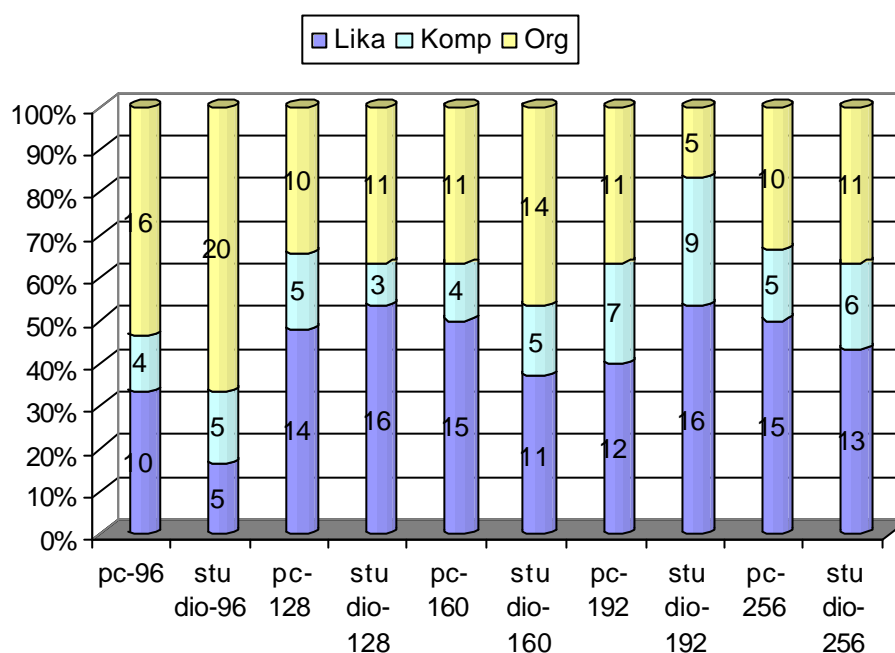


Diagram 6.8 Komprimerad musik på både pc och studio högtalare

6.12. Komprimerad klassisk musik

Vår studie visar en klar skillnad mellan systemen vid 96 kbps. Resultatet visar att majoriteten av respondenterna inte upplevde någon skillnad mellan original och komprimering på Pc:högtalarna och det omvända på studiomonitorerna. Vid övriga komprimeringsgrader är det ca 75% av respondenterna som uppfattar det komprimerade ljudet som lika bra eller bättre än originalet, se diagram ..

Våra resultat visar att det är svårare att höra skillnad mellan komprimerat ljud och original på klassisk musik jämfört med populärmusik. Detta tyder på att det antingen är så att respondenterna har svårare att uppfatta skillnaderna på denna kategori av musik vilket antagligen kan härledas till att dom inte är vana att lyssna på klassisk musik. Ett annat antagande som man kan göra är att de klassiska stycken som användes vid testet kanske inte hade ett lika brett frekvensomfång som populärmusikstyckena vilket gör att de inte drabbas i lika hög grad av den förlustgivande komprimeringen. I och med att vi fick så mycket som 40% som

upplevde det komprimerade ljudet som bättre än original (vid 160 kbps) kan det vara som så att komprimeringsalgoritmerna fungerade bättre för denna musik.

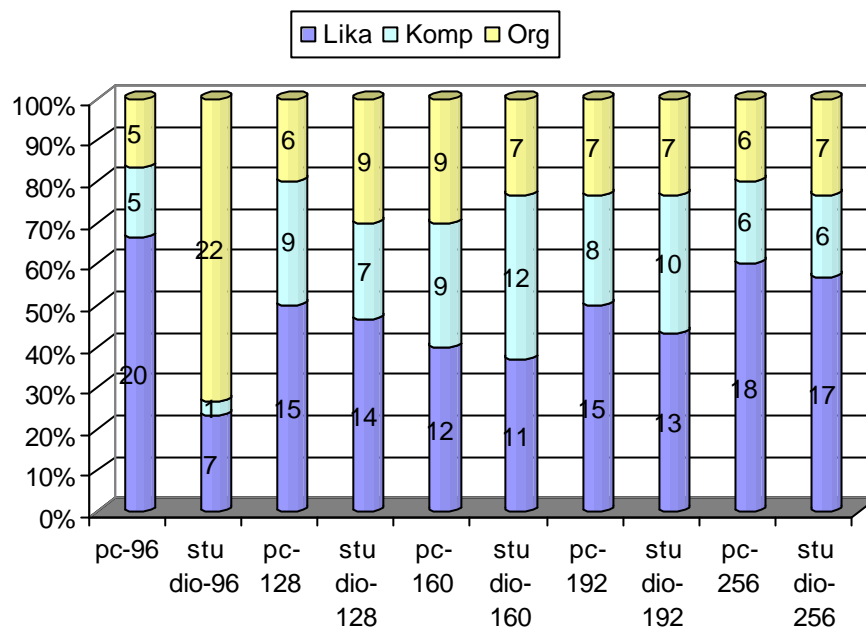


Diagram 6.9 Komprimerad klassisk musik på både pc och studio högtalare

7. Slutsats

De antagande som vi gör i vår slutsats skall absolut inte anses som generella utan endast som representativt för våra respondenter samt att de bygger på den inspelning och uppspelningsutrustning som vi har använt oss av i experimentutförandet.

Om man vid komprimerat tal slår ihop de respondenter som tycker att ljudparen låter lika och de som ansåg att det komprimerade ljudet var bättre får man fram att det är ca 80% som tycker detta med två undantag vilket är vid tal på studiomonitorer med 128 och 160 kbps komprimeringsgrad då ca 30% av respondenterna upplevde originalet som det bättre.

Resultatet av vårt experiment på populärmusik visar att en majoritet av respondenterna tycker att 96 kbps komprimering inte upplevs lika bra som originalet. När det gäller de övriga komprimeringsgraderna på pc:högtalarna så är det ca 65% av respondenterna som upplever det komprimerade ljudet som lika bra eller bättre. På studiomonitorerna så är det 192 kbps komprimering som upplevs som den bästa i och med att det är drygt 80% som anger detta som lika eller bättre än originalet. Detta resultat är ju intressant med tanke på att Fraunhofer hävdar att Cd-kvalitet uppnås vid 128 kbps.

När det gäller klassisk musik på pc:högtalare så hör de flesta ingen skillnad mellan original och komprimering. Detta antagande bygger på att det är så mycket som 2/3 av respondenterna som anser att det låter lika vid en jämförelse mellan 96 kbps komprimering och original samt att vi genom hela testet har endast fått mellan 15-30 % som har upplevt originalet som det bättre ljudet.

Våra respondenters upplevelser av experimentet har lett fram till följande antagande.

- Tal som skall spelas upp på pc:högtalare kan komprimeras med valfri bithastighet ända ned till 96 kbps.
- Tal för uppspelning på bättre ljudsystem kan komprimeras på liknande sätt men för största möjliga ljudkvalité bör en komprimeringsgrad över 160 kbps väljas.
- Populärmusik för användning på pc:högtalare bör komprimeras med minst 128 kbps.
- Populärmusik för uppspelning på bättre ljudanläggningar bör komprimeras i 192 kbps (vilket fick bättre resultat än 256 kbps).
- Klassisk musik på pc:högtalare kan komprimeras ända ned till 96 kbps.
- Klassisk musik för uppspelning på stereoanläggningar bör komprimeras med minst 128 kbps.

8. Referenser

- Andersson, Bibbi & Sönnebo, Lars (1998).
<http://www.tpd.se/talbocker/utredning/u27.htm> –hämtat 2001-11-14
- Andersson, Daniel (2001), *Daniel lyssnar just nu*
<http://daniel.lyssnar.just.nu/> -hämtat 2001-11-14
- Arlinger, Stig, & Lidén, Gunnar (1985). Akustik och psykoakustik. G. Lidén (red.),
Audiologi, Almqvist & Wiksell Förlag AB, Stockholm,
- Backman Jarl (1998), *Rapporter och uppsatser*, Studentlitteratur, Lund
- Bodelius, Per.(1998), *Director – grunder i multimediaproduktion* – Liber Ekonomi,
Malmö
- Denscombe Martin (2000), *Forskningshandboken* (P. Larson övers),
Studentlitteratur, Lund
- DePoy, Elizabeth & Gitlin, Laura N (1999), *Forskning- en introduktion*,
Studentlitteratur, Lund
- Ejlertsson, Göran (1996), *Enkäten i praktiken - en handbok i enkätmetodik* ,
Studentlitteratur, Lund
- Eriksson, Mats (2001) *Ljudkompression och audio på webben*
http://www.sit.fi/~meerikss/nywebb/Ljudkomprimering.html#_Toc507997145 –
hämtat 2001-12-20
- Fraunhofer (2000). *MPEG Audio Layer-3* -Institut Intergrirte Schaltungen
<http://www.iis.fhg.de/amm/techinf/layer3/index.html> –hämtat 2001-11-13
- Hacker, Scot (2000), *MP3- The Definite Guide*, O'Reilly & Associates, Inc.,
Sebastopol
- Hägerby, Birgitta (1971), *Teknisk Audiologi*, Almqvist & Wiksell Förlag AB,
Stockholm
- Jansson, Dan.(2000). *Whats is MP3?* – DJ-Media.com
<http://www.dj-media.com/> -hämtat 2001-11-14
- Kindersley, Dorling (1997), *Allt om Multimedia – En utförlig guide* (S. Kjell
övers), Bokförlaget Bonnier Alba (Originalarbete publicerat 1996)
- Mediaspec (2001) *The Genes of Genelec*
<http://www.mediaspec.co.uk/site2000/wax/genesofgenelec/genelec.htm> -hämtat
2001-12-11
- Mülen, Hans (1998). *Digitalt ljud*
<http://internet.physto.se/fil/ljud.html> -hämtat 2001-11-13
- Nielsen, Åge (2001) Kodning, komprimering, reducering & lagring av Digitaliserat
ljud, *Dator Magazin* , 116-127.
- Nilsson, Marianne.(1998). *Tidsbaserade medier*
<http://www.dis.uu.se/~mrytther/tbm/> -hämtat 2001-11-15
- Nilsson, Marianne.(1998b). *Introduktion till Digitalt ljud*
<http://www.dis.uu.se/~mrytther/tbm/introtillljud.html> -hämtat 2001-11-16

Nyquist, Bjarne & Fink, George (1995) *Tekniken i musiken*, KMH Förlaget, Sundbyberg

Patel, Runa & Davidson Bo (1994), *Forskningsmetodikens grunder*, Studentlitteratur. Lund

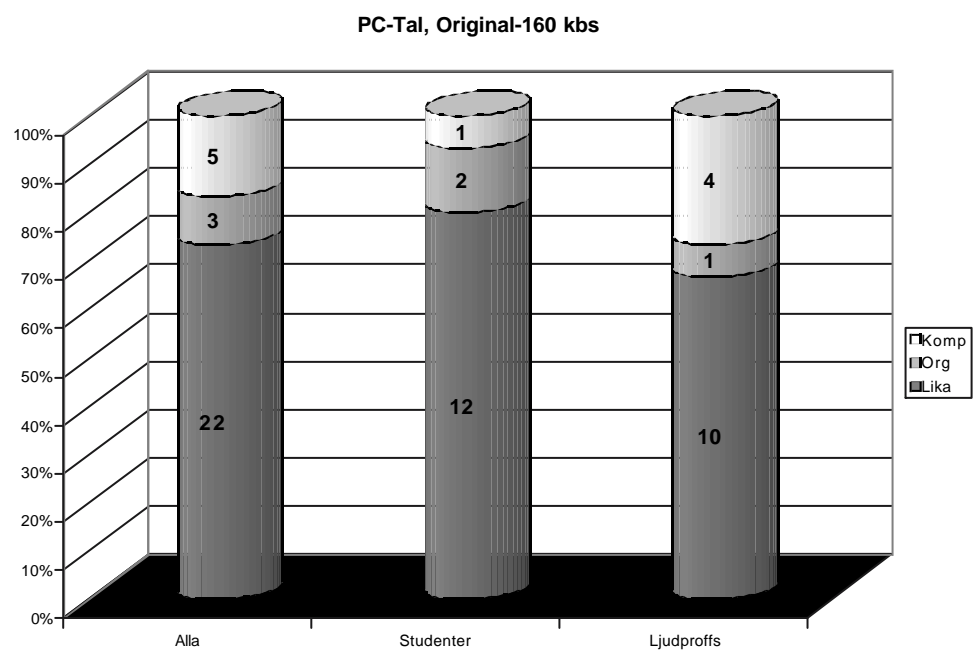
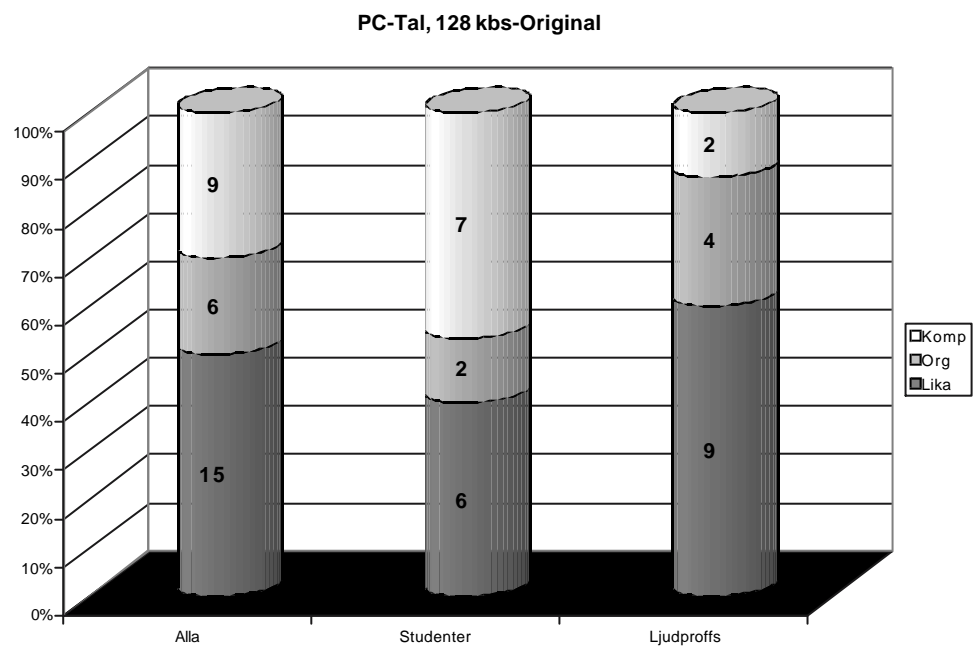
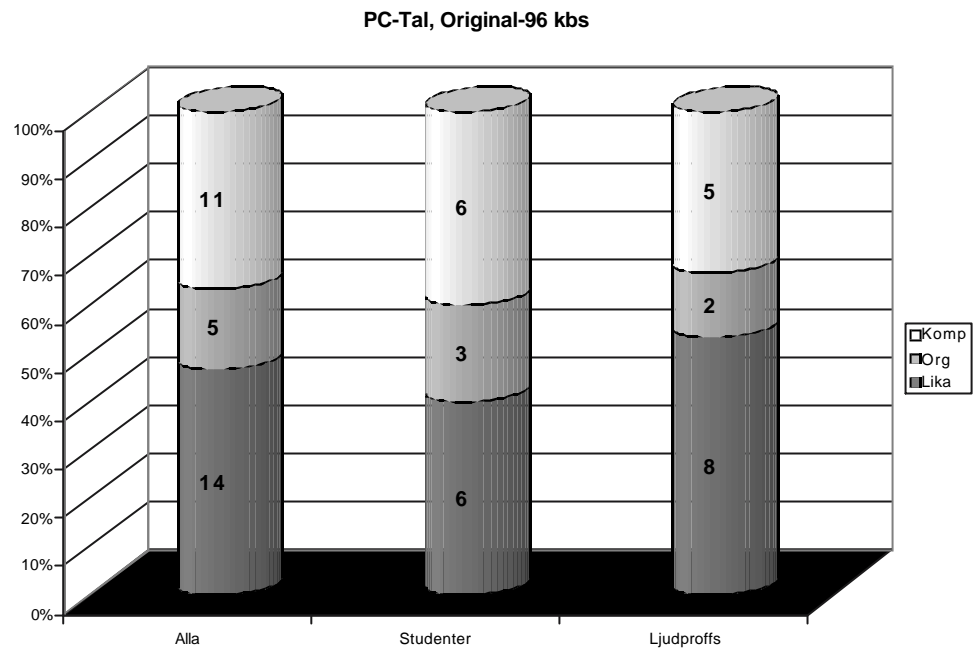
Patel, Runa & Tebelius, Ulla (1987) *Grundbok i forskningsmetodik*, Studentlitteratur, Lund

Smelink (2001), *Ordlista för arbetsmiljö*

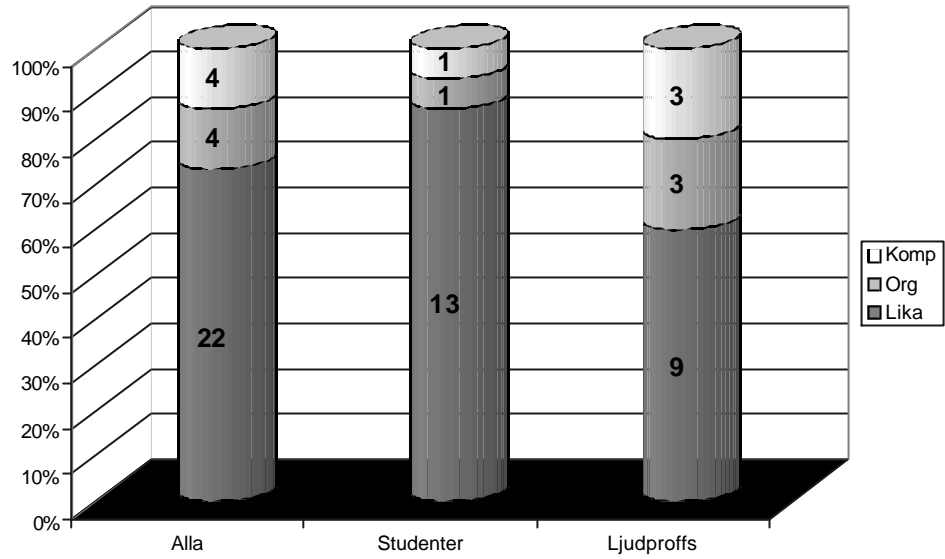
<http://www2.smelink.se/startadriwa/miljokval/arbetsmiljo/alag/aord/l.htm> -hämtat 2001-11-14

Whatis?com (2000). *Multimedia & Graphics Categories*

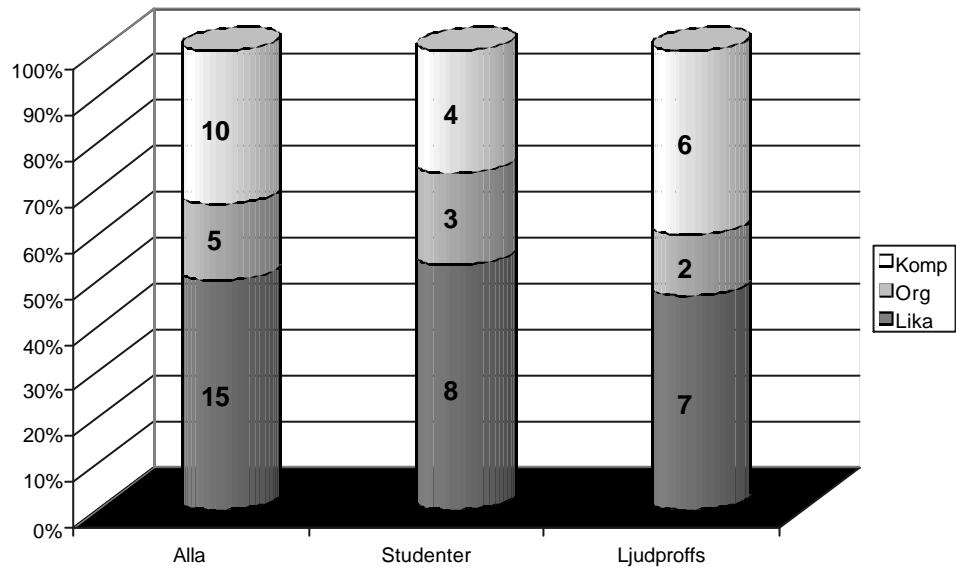
http://whatis.techtarget.com/WhatIs_Definition_Page/0,4152,213472,00.html – hämtat 2001-11-14



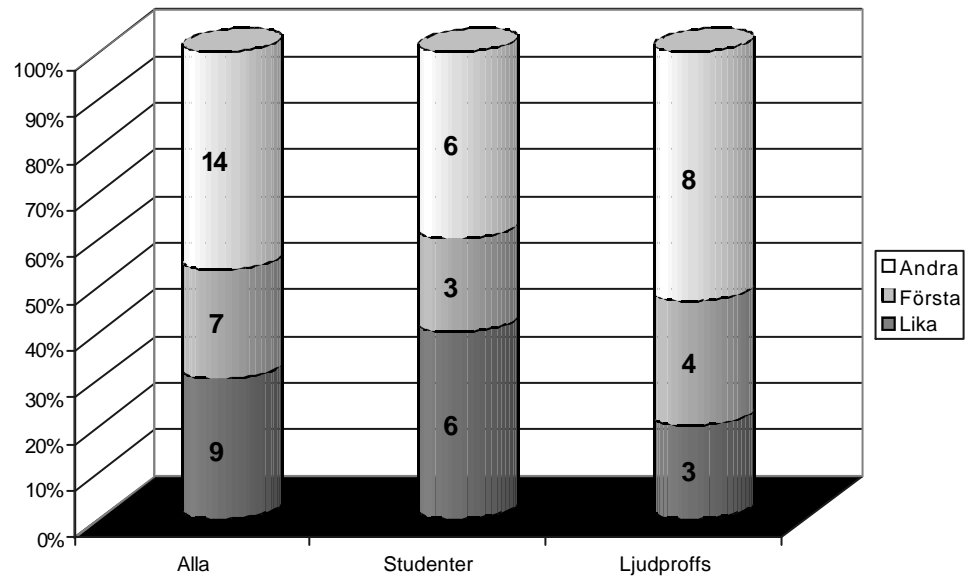
PC-Tal, Original-192 kbs



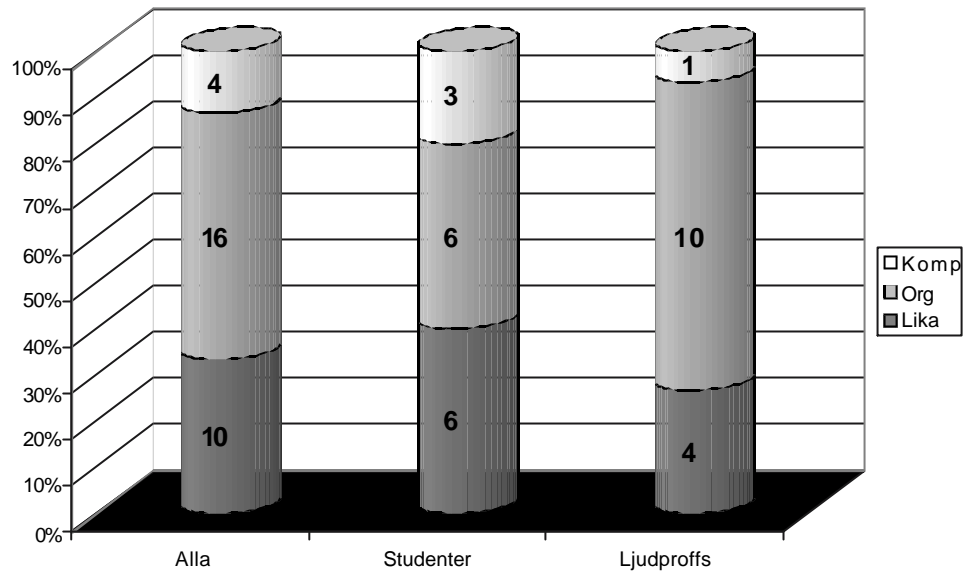
PC-Tal, 256 kbs-Original



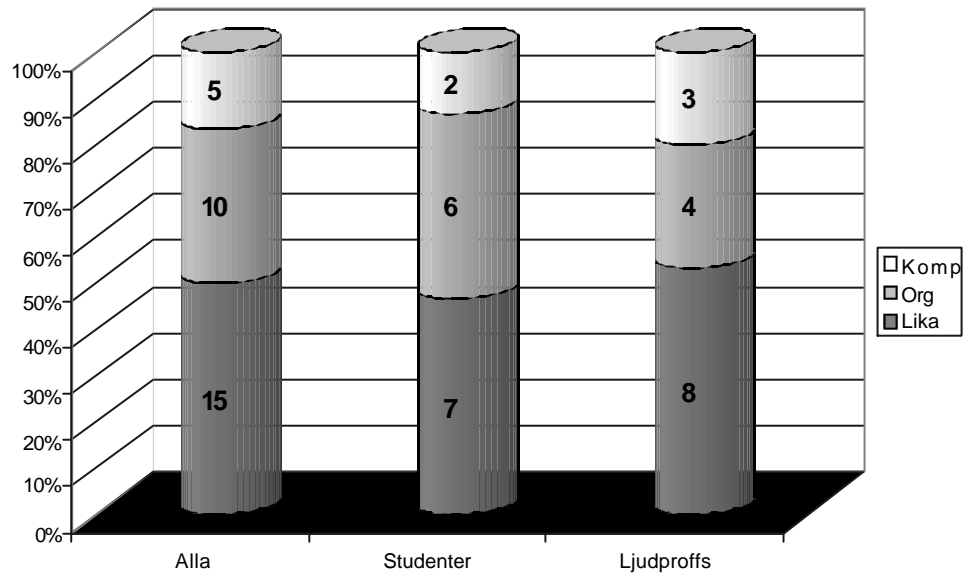
PC-Tal, Original-Original



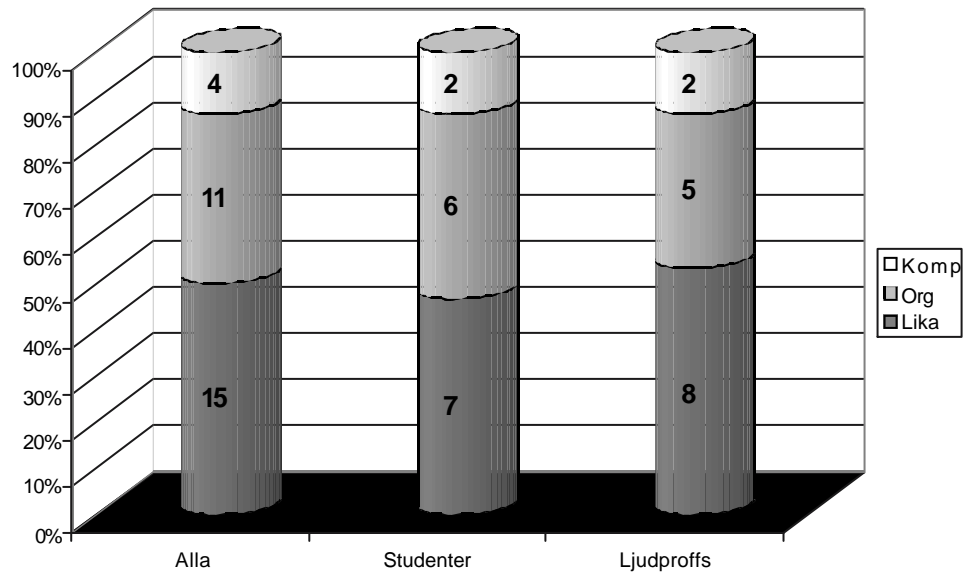
Pcmusik 96kbs-Orig



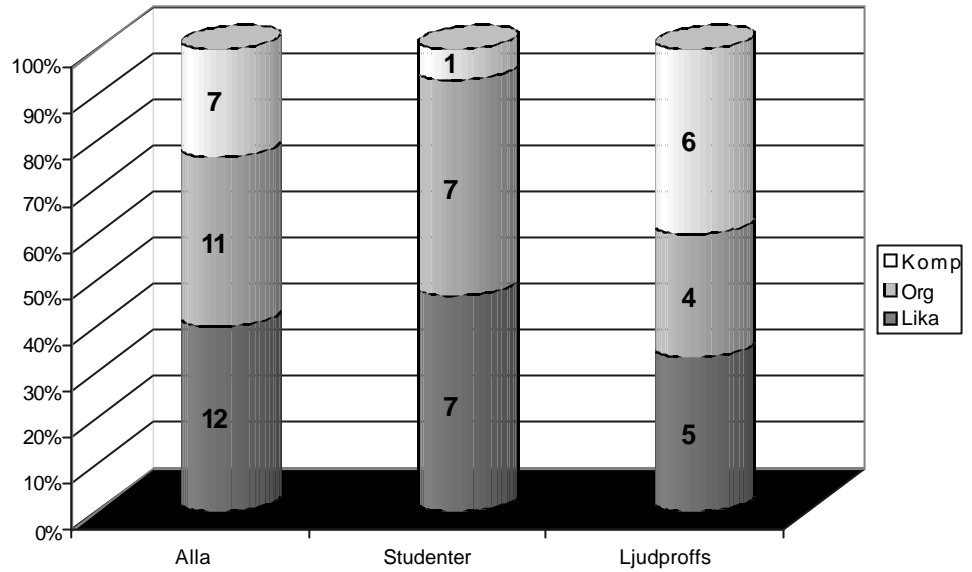
Pcmusik, Orig-128 kbs



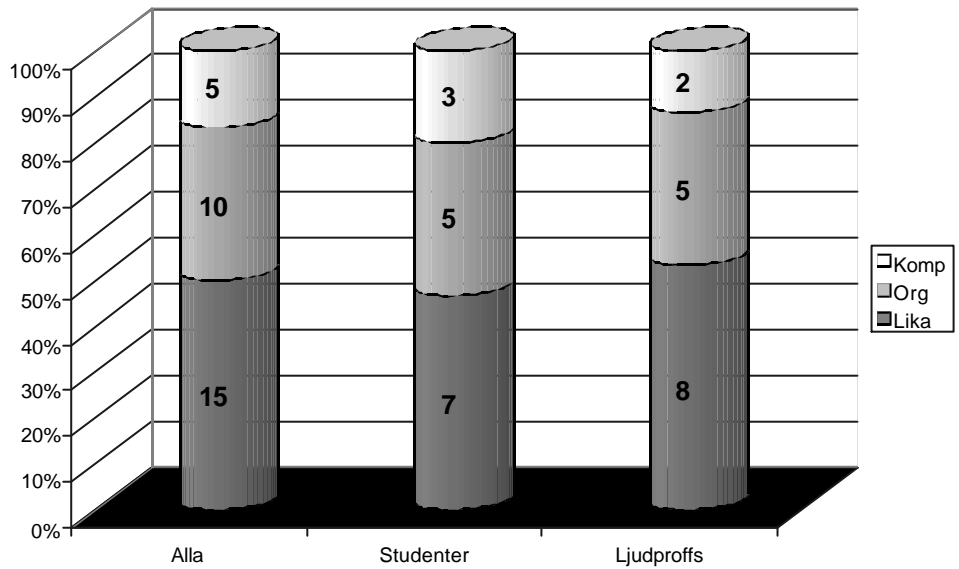
Pcmusik, orig-160kbs



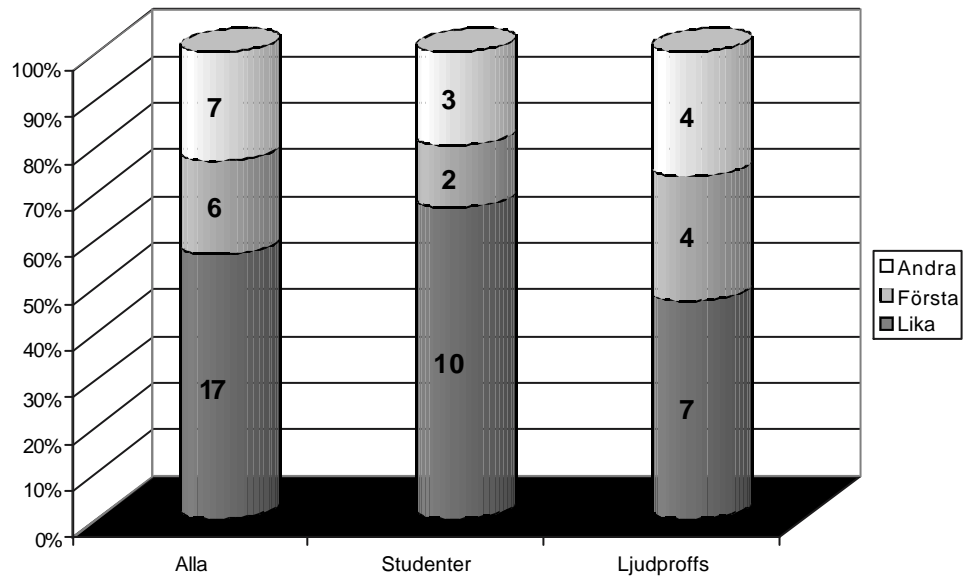
Pcmusik, Orig-192kbs



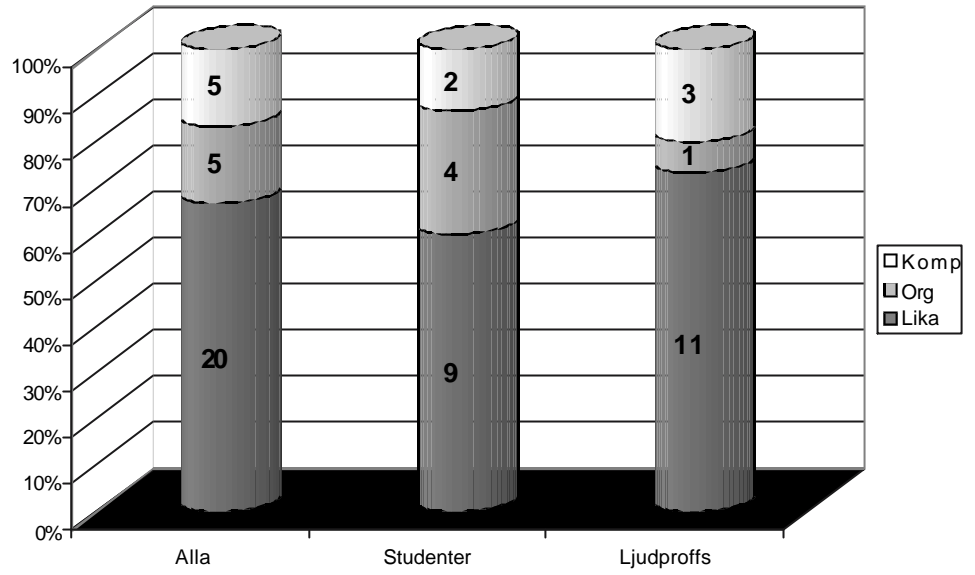
Pcmusik, 256kbs-Orig



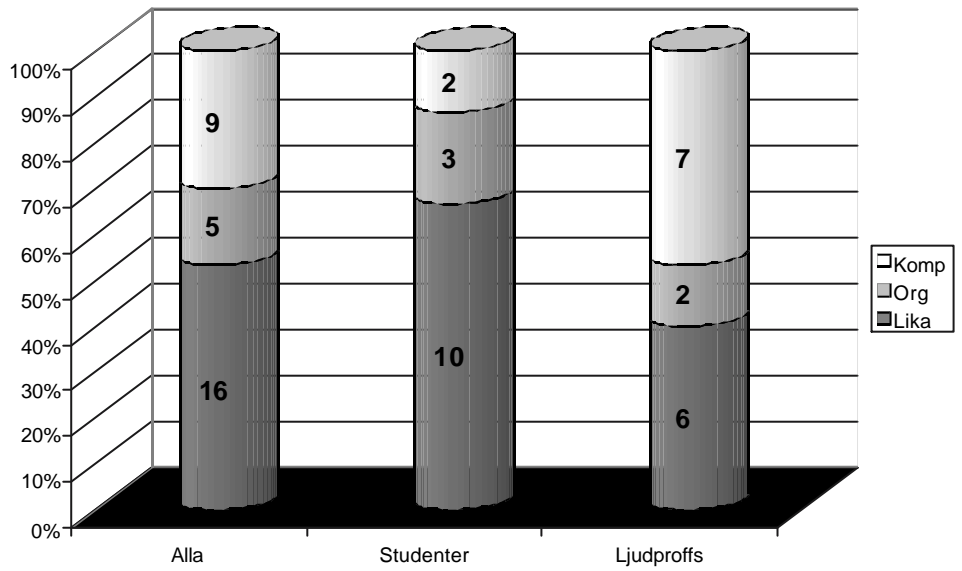
Pcmusik, Orig-Orig



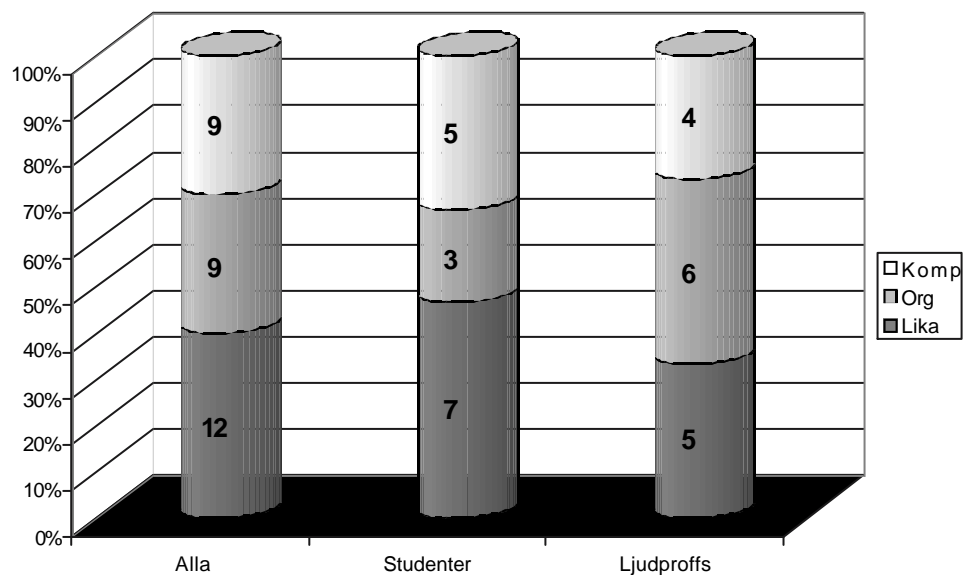
PC-klass, Orig-96 kbs



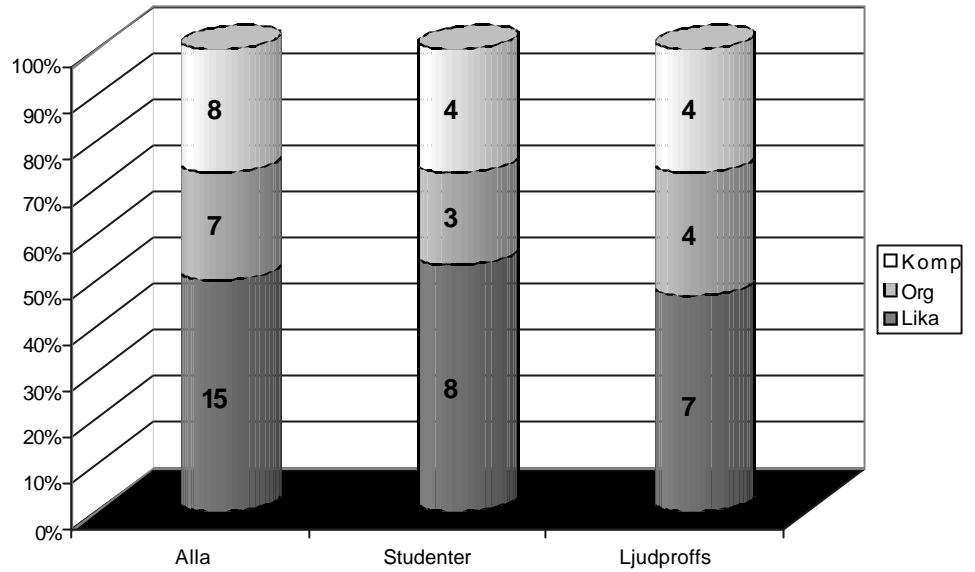
PC-klas, 128 kbs-Orig



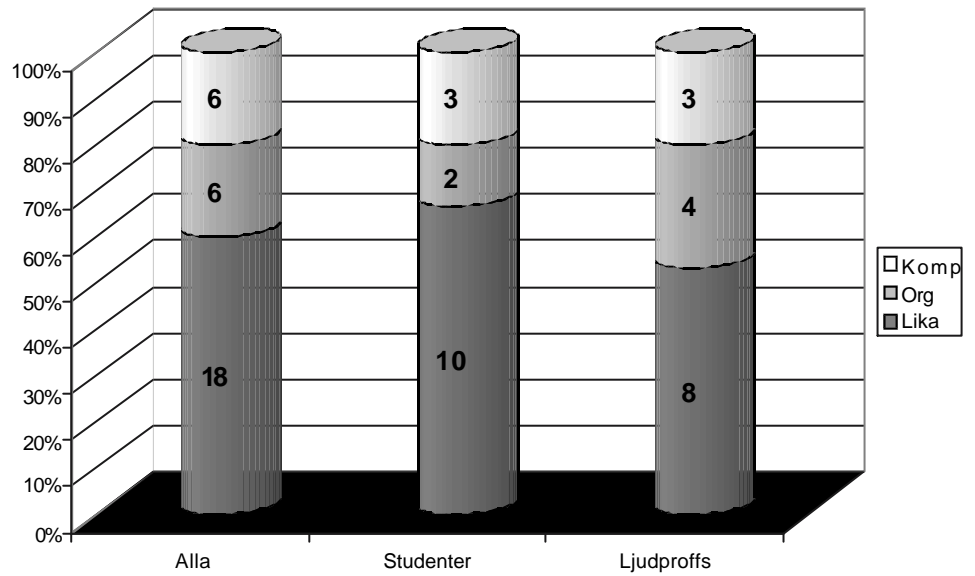
PC-klas, 160 kbs-Orig



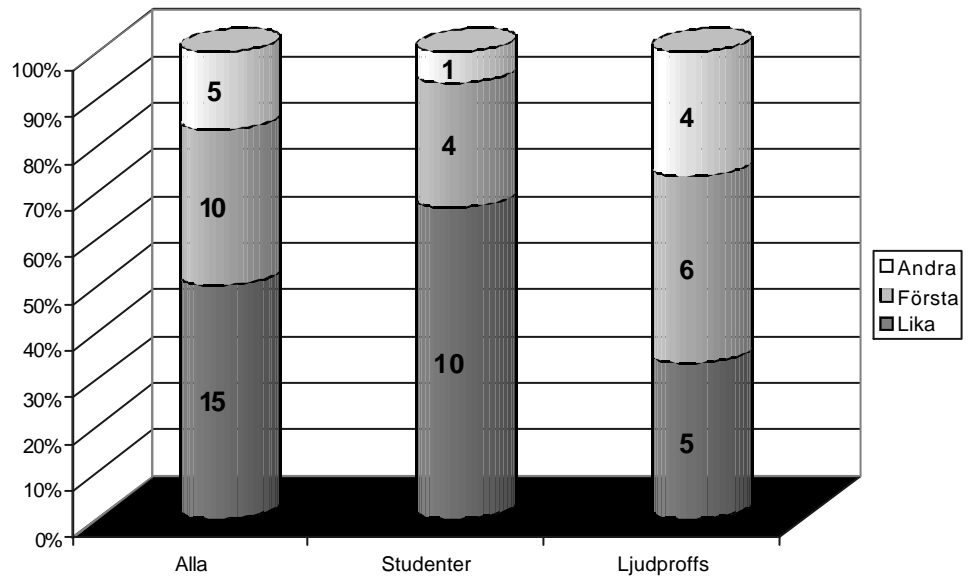
PC-klas Orig-192kbs



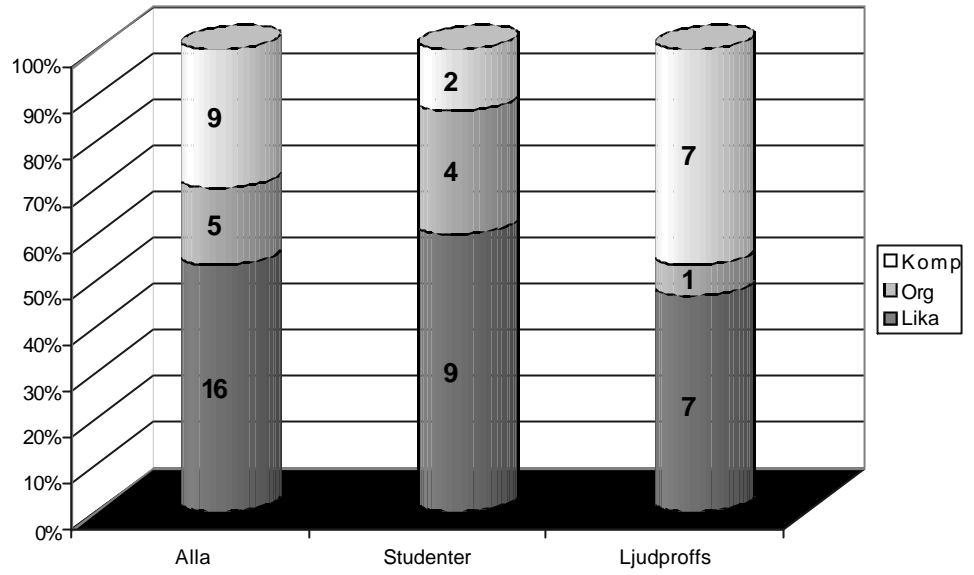
PC-klas, 256kbs-org



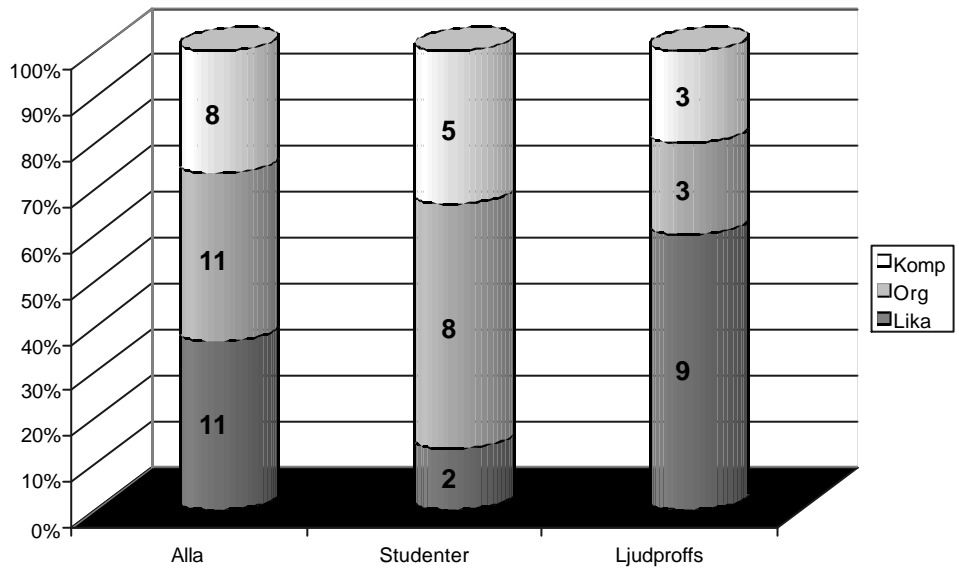
PC-klas Orig-Orig



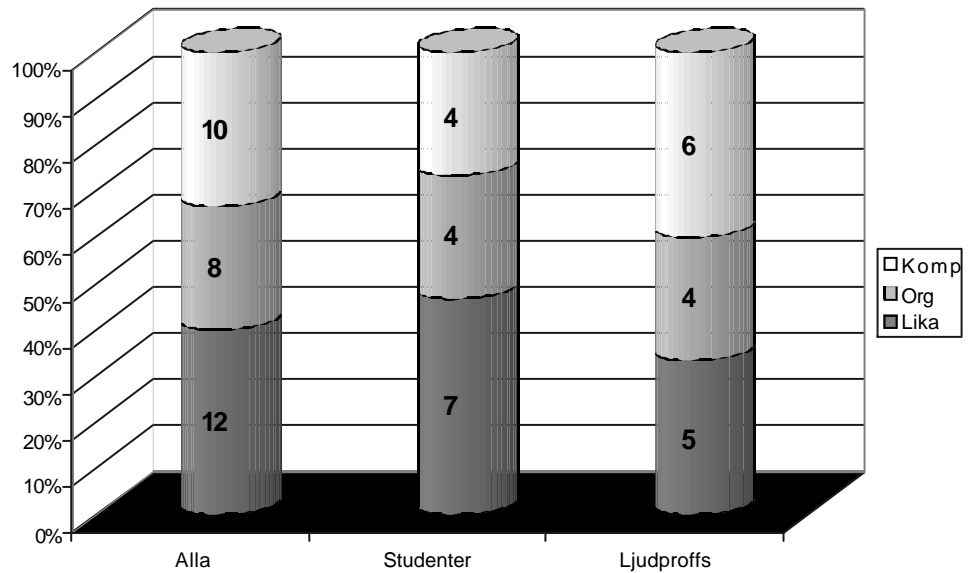
Studiotal org-96kbs



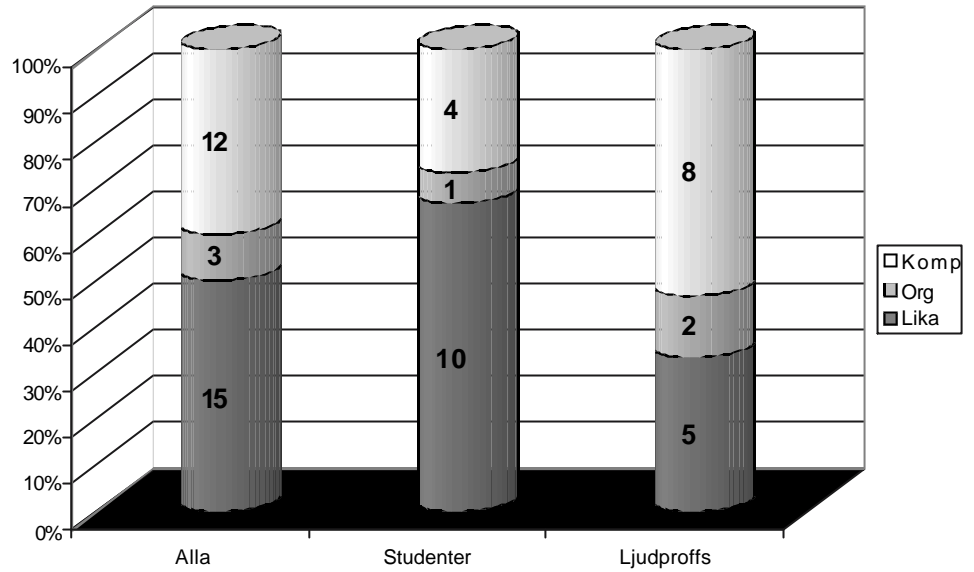
Studiotal 128kbs-org



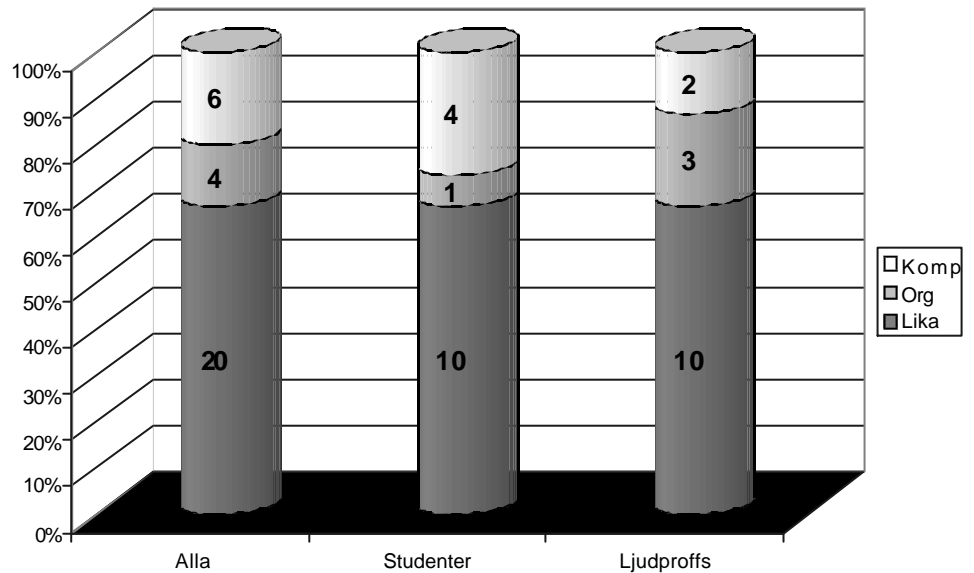
Studiotal org-160 kbs



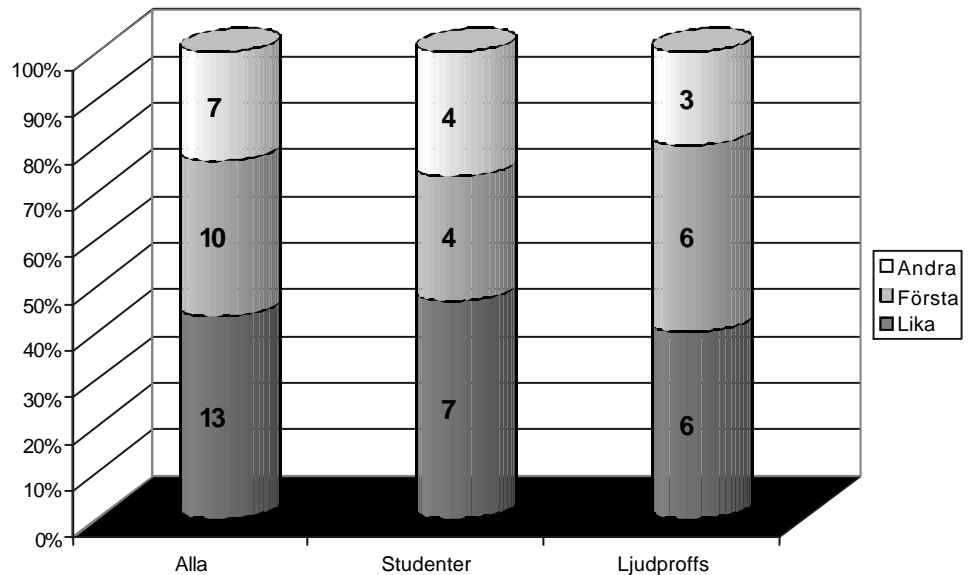
Studiotal 192 kbs-org



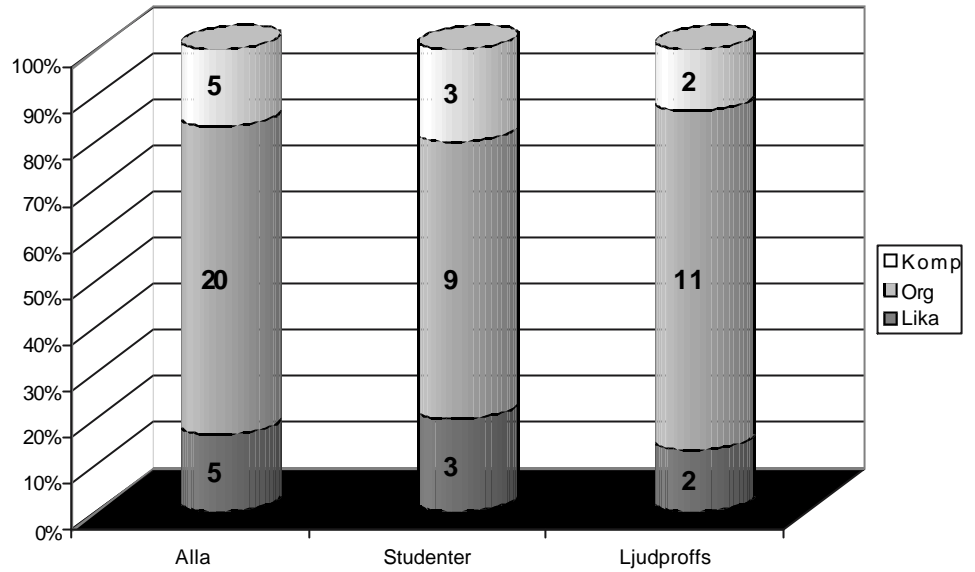
Studiotal 256kbs-Org



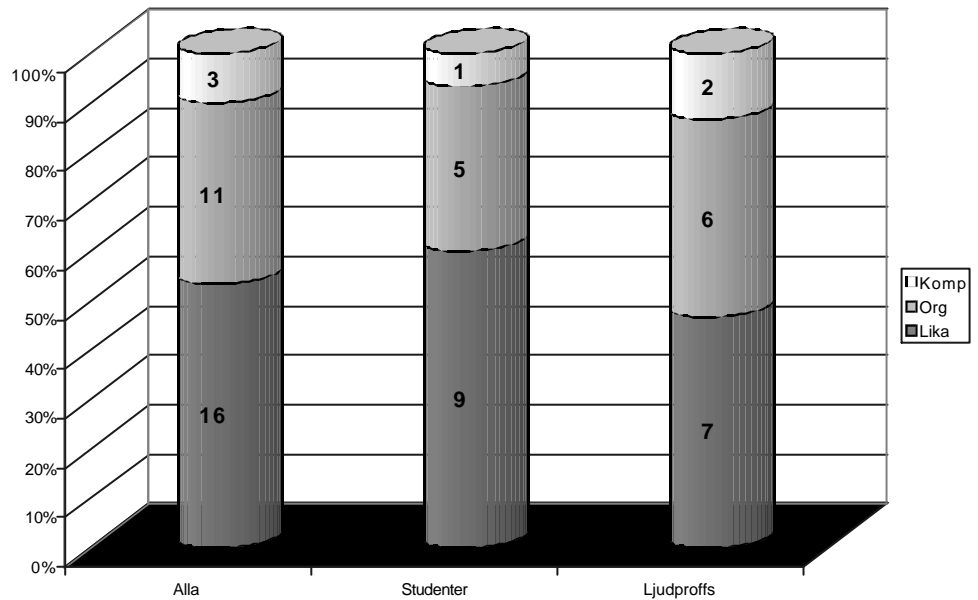
Studiotal Org-Org



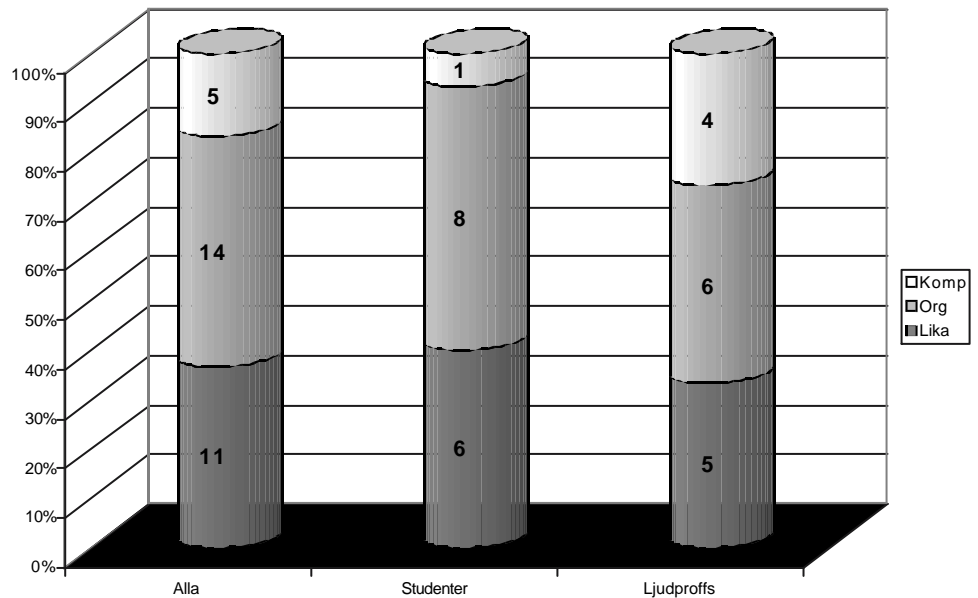
Studiosus org-96kbs



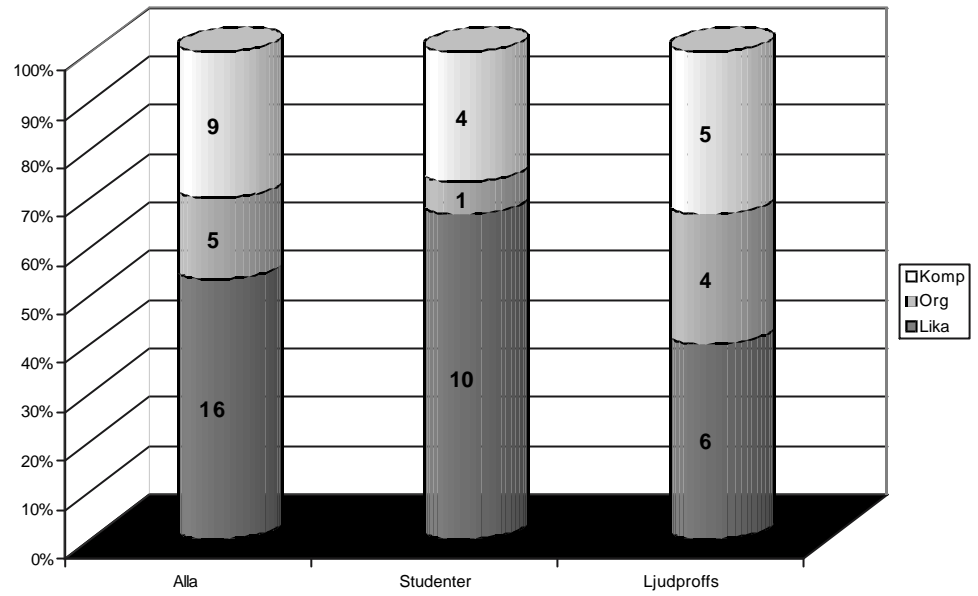
Studiosus 128kbs-org



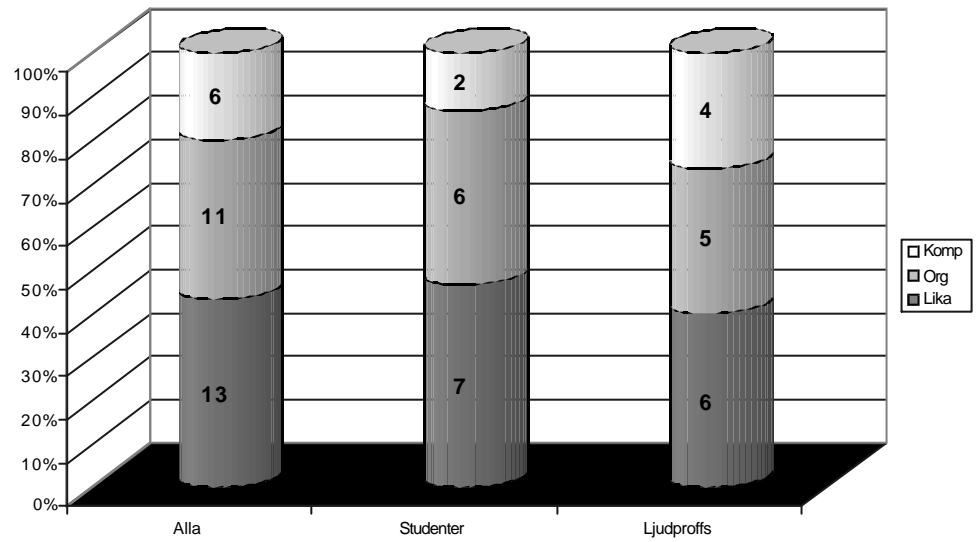
Studiosus 160kbs-org



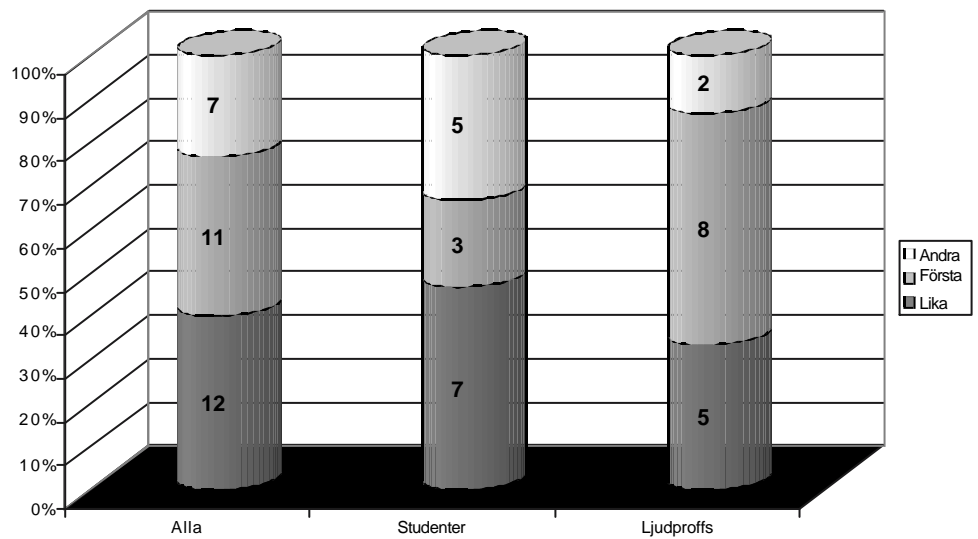
Studiomus org-192kbs



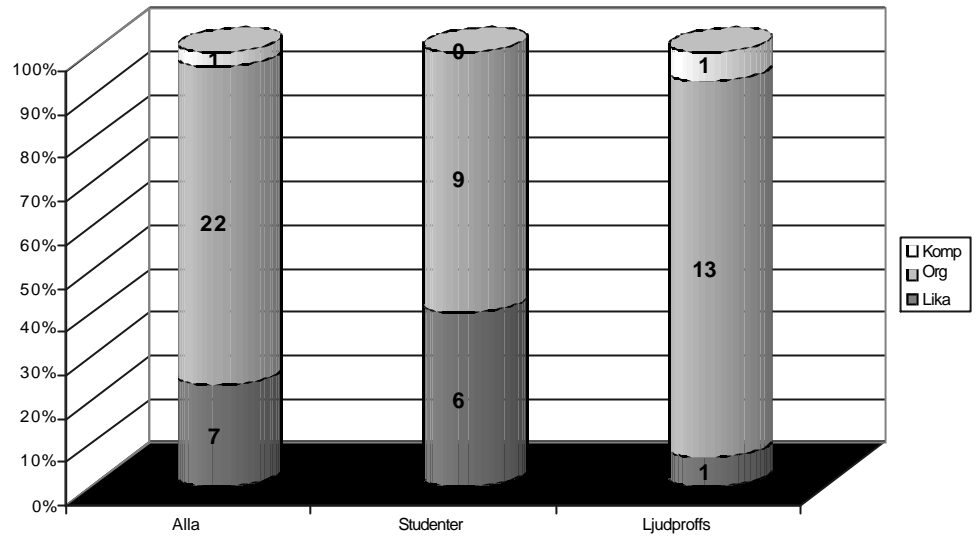
Studiomus 256kbs-org



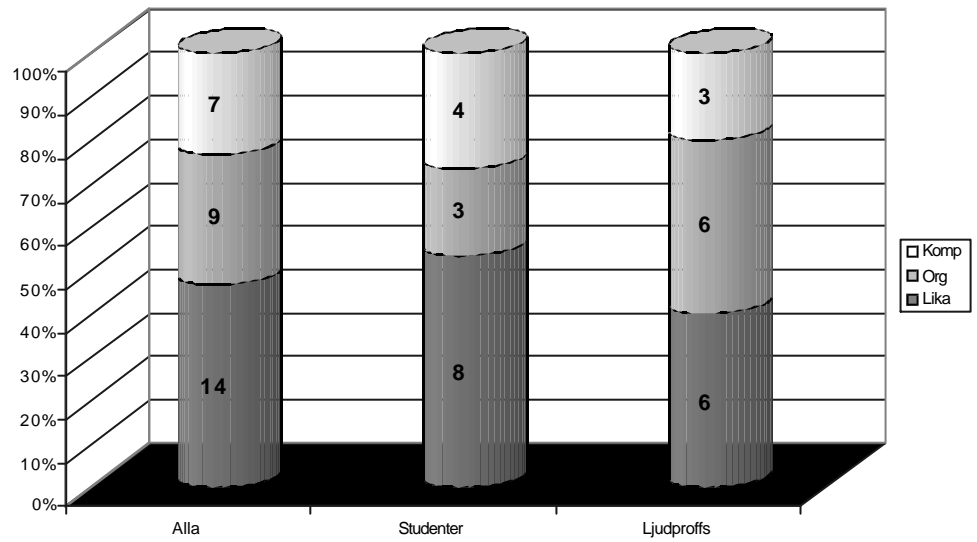
Studiomus org-org



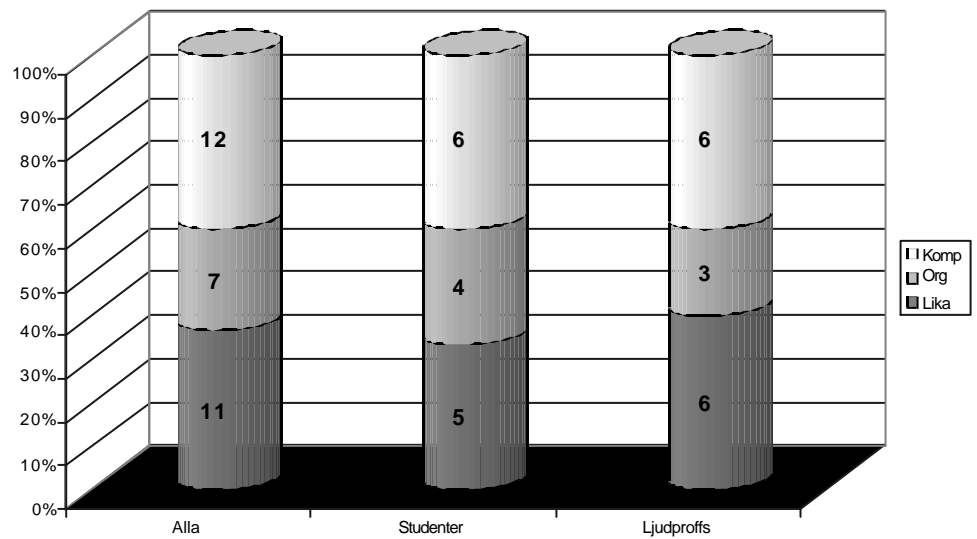
Studioklas 96kbs-org



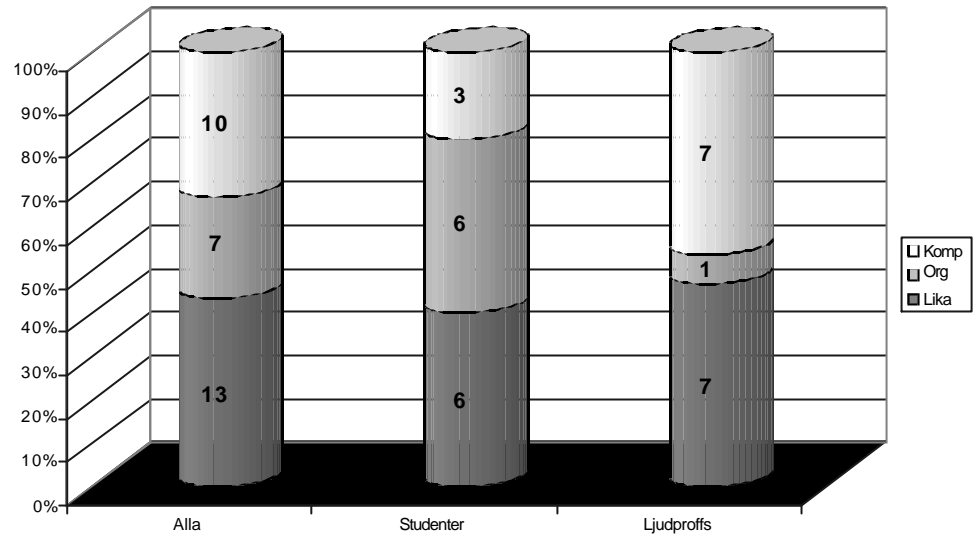
Studioklas org-128kbs



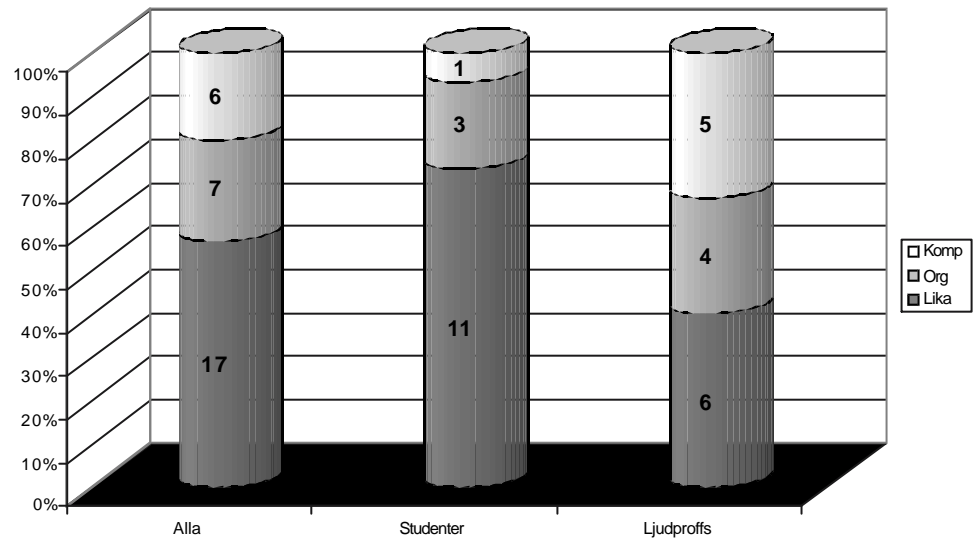
Studioklas org-160kbs



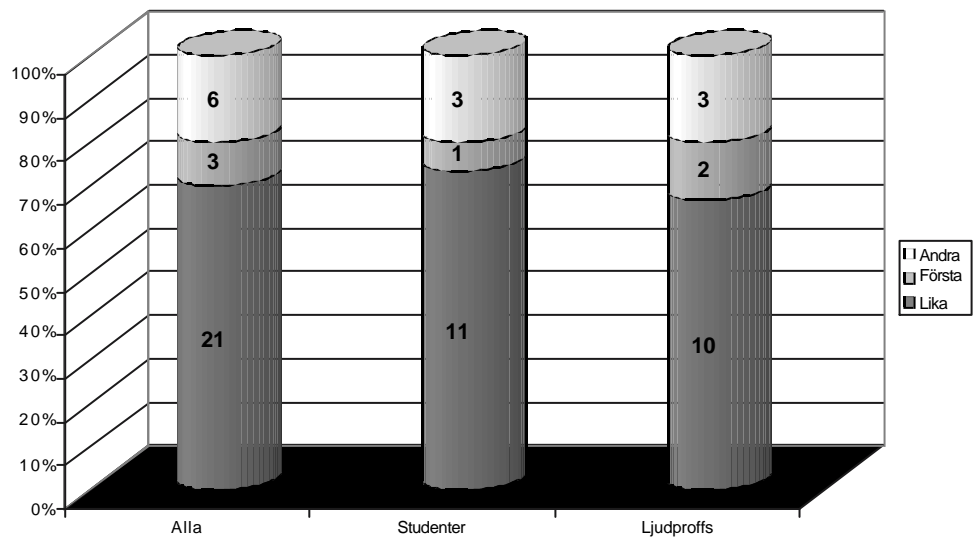
Studioklas org-192kbs



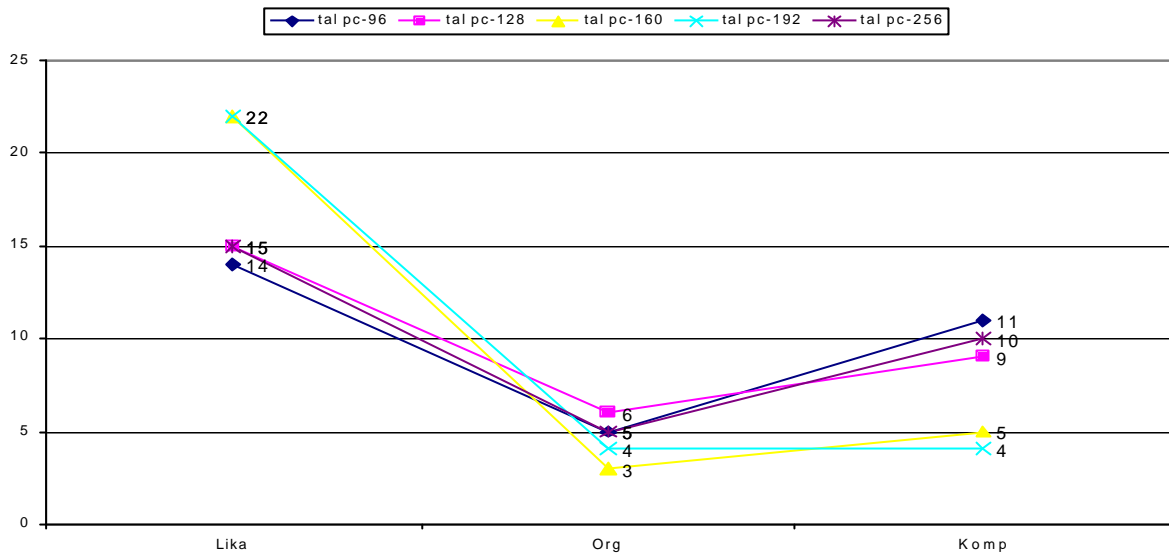
Studioklas 256kbs-org



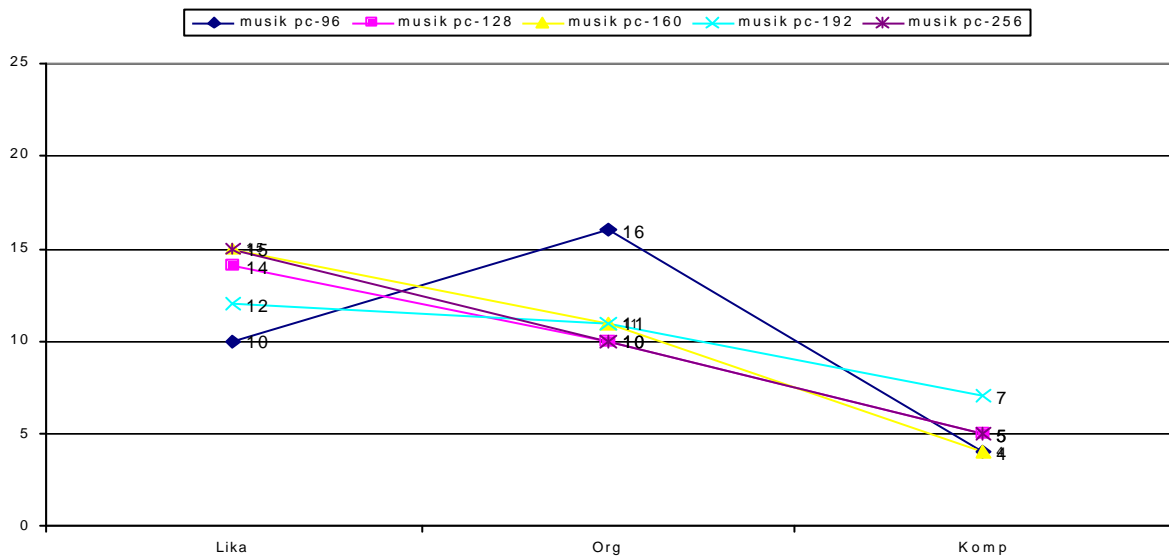
Studioklas org-org



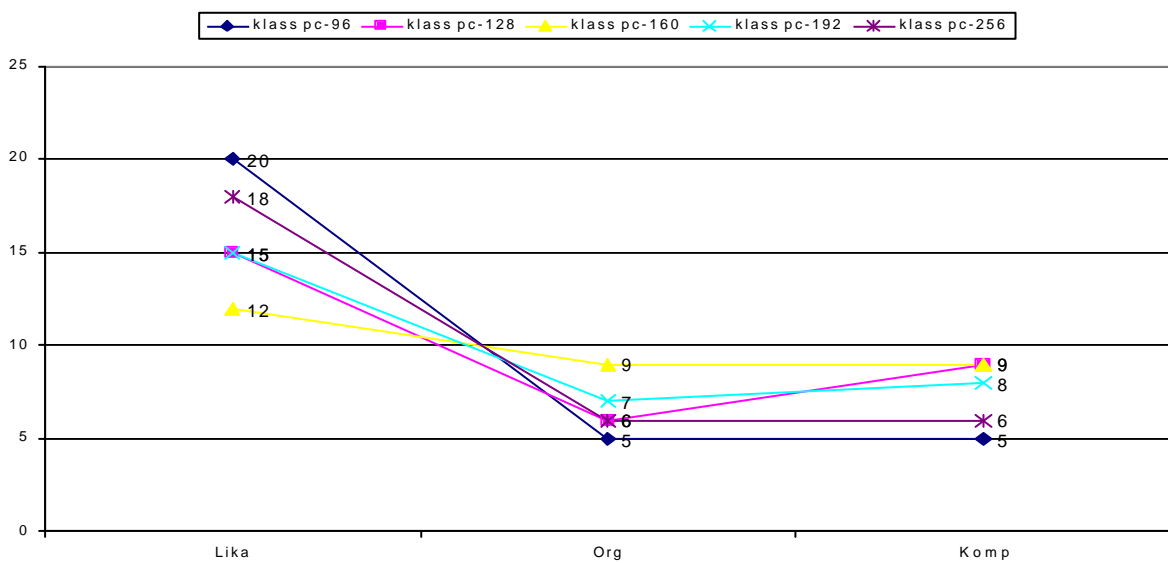
Tal på PC-högtalare



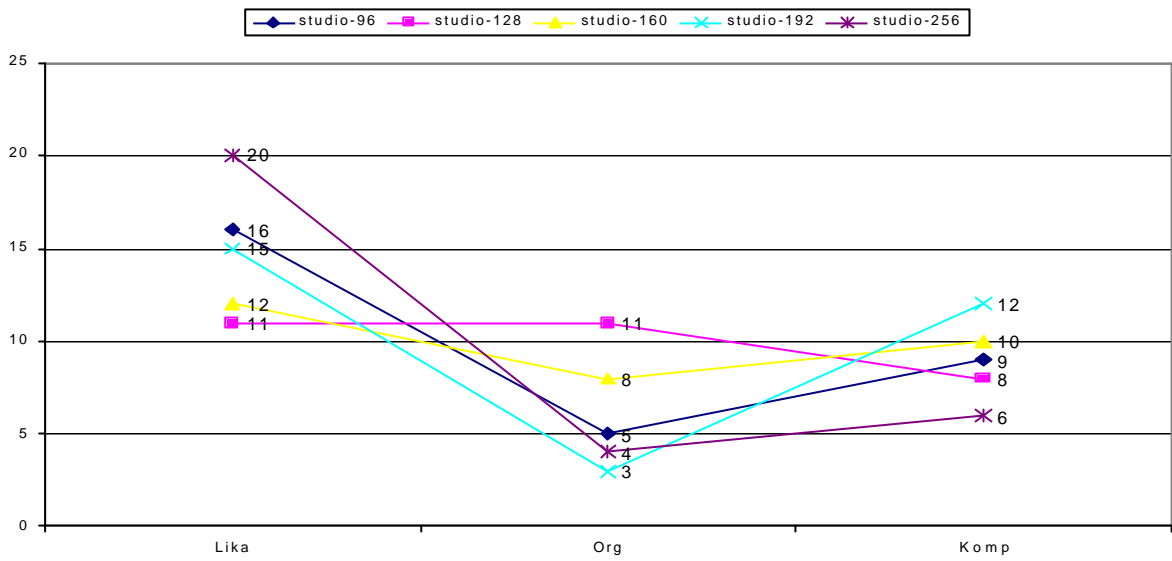
Populärmusik på PC-högtalare



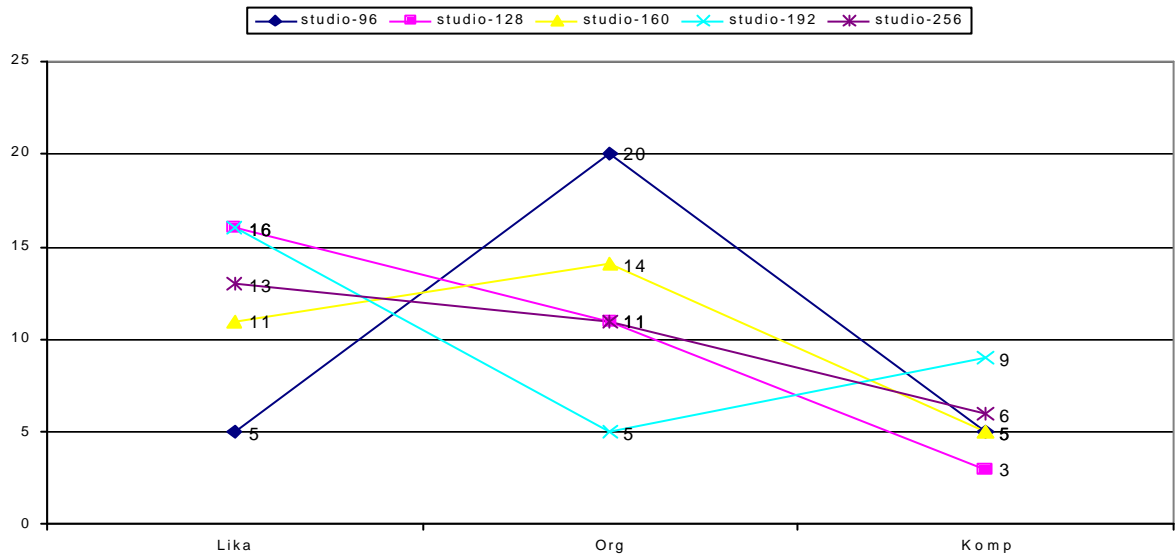
Klassisk musik på PC-högtalare



Tal på studiomonitorer



Populärmusik på studiomonitorer



Klassisk musik på studiomonitorer

