



*Skrifter från*

Lyssnande Lund  
Ljudmiljöcentrum vid Lunds  
universitet

*Rapport nr 3*

# OPERATIVA LJUD

Publisher: Sound Environment Centre, Lund University

Text © Editorial, Frans Mossberg; individual chapters, the contributors, 2006

This book is licensed under an Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License. This license allows users to download and share the article for non-commercial purposes, so long as the article is reproduced in the whole without changes, and the original authorship is acknowledged. (See full terms and conditions here:<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

e-ISBN 978-91-89415-37-9

DOI <https://doi.org/10.37852/oblu.150>

ISSN 1653-9354

*Publications from the Sound Environment Centre at Lund University Report no. 3*

Printed in Sweden by Media-Tryck, Lund University, Lund 2006

(Citation, e.g.: Mossberg, F. (Red.) (2006). Operativa ljud. (Skriftserie; Vol. 3). Ljudmiljöcentrum vid Lunds universitet, DOI: <https://doi.org/10.37852/oblu.150>)

Information about the Sound Environment Centre, Lund University, can be found here: <https://www.lmc.lu.se/>



**Skrifter från**  
**Lyssnande Lund - Ljudmiljöcentrum vid Lunds universitet**

**Rapport nr 3**

**Operativa ljud**

*Texter från seminarium den 22 september 2006  
arrangerat av  
Lyssnande Lund – Ljudmiljöcentrum vid Lunds universitet*

---

Lund 2006

ISSN 1653-9354

---

Red: Frans Mossberg, Lund 2006  
Tryck: Media-Tryck, Lund 2006

# Innehåll

Inledning	1
Pernilla Ulfvengren <i>Naturliga varningsljud i människa-maskinsystem</i>	3
Marcus Sanchez Svensson <i>När ljud pockar på uppmärksamhet</i>	13
Dan Gärdenfors <i>Playing with auditory interfaces</i>	25
Petter Alexanderson, Konrad Tollmar <i>Operational sounds – The factory soundscape as carrier of information</i>	36
Kirsten Rasmus-Gröhn <i>Icke-visuella audio-haptiska datorgränssnitt</i>	45



## Inledning

Denna tredje utgåva i Lyssnande Lund – Ljudmiljöcentrums rapportserie är tillkommen i samband med ett tvärvetenskapligt seminarium den 22 september 2006 med *Operativa ljud* som tema. Seminariet behandlade skilda aspekter på ljud som informationsbärare och hur man kan använda operativa ljud för att synliggöra egenskaper och funktioner, samt för att ge återkoppling på olika handlanden.

Kvaliteten på maskinsystemens visuella och auditiva gränssnitt i interaktionen med människan kan många gånger vara livsavgörande. Under rubriken *Naturliga varningsljud i människa-maskinsystem* diskuterar *Pernilla Ulfvengren* varningssystem i flygplan och hur dessa system kan förbättras. Efter Gottröra-olyckan t.ex. visade piloternas erfarenheter att de i realiteten upplevde delar av varningssystemen som både otillräckliga och förvirrande. Genom experimenterande simulationer i nära samarbete med piloter och flygpersonal har Ulfvengren sökt finna ut hur auditiva larmsystem kan optimeras.

Ljud har ofta på liknande sätt som telefonsignaler funktionen av att vara "tillkallare" som kräver svar och reaktion. *Marcus Sanchez Svensson* granskar hur larmsystem på en dialysavdelning fungerar och personalens reaktioner på detta. Han lägger tyngdpunkten på ljud som ett socialt fenomen och frågar sig vad som händer när ljud pockar på uppmärksamhet. Med detta lägger han också fokus på de sociala överenskommelserna bakom ett larms funktion som informationsbärare. I likhet med Ulfvengren ser han faran i att ogenomtänkta ljudresurser kan hindra koordinationsarbete stället för att stödja det. En av framtidens stora utmaningar, skriver han, är att undvika ett "inferno av datorer som ständigt pockar på vår uppmärksamhet med signaler och information om det ena och det andra".

*Dan Gärdenfors* har utvecklat dataspel för blinda och handikappade där auditiva gränssnitt utnyttjas för interaktion mellan användare och maskin. Den övervägande delen digitala media för synskadade använder sig av text, inspelade röster eller talsyntes. Gärdenfors frågar sig om det är möjligt att tillverka spel för denna användargrupp med ljud och musik som ersättare för grafiken. Genom att



utnyttja grafik med starka kontraster kan en synskadad spelare orientera sig och agera med hjälp av auditiva signaler och responser. Ljudens tendenser att maskera varandra på andra sätt än simultant uppfattade komponenter i synbilder, innebär dock problem som man arbetar med. Gärdenfors menar i en avslutande observation att både ljuddesigners och spelutvecklare står i behov av mer forskning runt ljuds betydelser, en "auditiv ikonografi".

*Petter Alexanderson* och *Konrad Tollmar* behandlar ljudlandskap i industri och verkstad. De tar utgångspunkt i ljudmiljön i en kemisk fabrik och i hur operatörer tillskriver innebörder till ljuden i sitt dagliga arbete. Maskinernas ljud bär information om deras status och verksamhet, så att operatörerna med hjälp av dessa kan identifiera om något inte fungerar som det skall. Genom att bygga upp kartotek av inspelningar från specifika punkter i industrilandskapet kan kunskaperna ökas om maskinparkens funktionalitet och med detta även möjligheterna att vidta lämpliga åtgärder vid driftsstörningar. Författarnas bidrag pekar på ökande behov av "medvetandegörande service" (awareness services) i samhället.

Slutligen ger *Kirre Rasmus-Gröhn* exempel på forskning som föregår vid *Certec* vid Lunds universitet om innovativa datorgränssnitt för handikappade. Där använder man sig bl.a. av ljud i kombination med känsel-återkoppling i sina applikationer. Ett exempel på projekt är ett audio-haptiskt ritprogram där pennans position kan höras genom tonändring i en ljudsignal, ett annat är en audio-haptisk virtuell trafikmiljö som ger synskadade möjlighet att simulera trafiksituationer innan han eller hon går ut i "den riktiga" miljön.

Ljudmiljöcentrum tackar samtliga författare och medverkande för deras bidrag.

Lund den 11 oktober 2006

Frans Mossberg  
Koordinator/producent  
Lyssnande Lund  
Ljudmiljöcentrum vid Lunds universitet

# Naturliga varningsljud i människa-maskinsystem

Pernilla Ulfvengren  
Industriell arbetsvetenskap  
INDEK KTH

pernilla.ulfvengren@indek.kth.se

## Inledning

Utan tvekan har ljud och tystnad ett stort inflytande på vår tillvaro. Utvecklingen av dagens samhälle har i mångt och mycket berövat oss från många vardagliga användbara ljud. Samtidigt har många nya ljud tillförts, som inte alla gånger är så användbara som de skulle kunna vara.

Vårt hörselsystem är inte uppbyggt för att passivt ta emot ljud, utan för att aktivt lyssna och söka information i vår omgivning. Att lyssna efter plötsliga förändringar eller andras varningsrop är en viktig funktion och det faktum att naturen inte försett oss med öronlock talar för att just hörseln ska vara en öppen kanal, ständigt lyssnande. Därför är det fullständigt naturligt att lägga till och använda skapade ljud i tekniska system som varningar och med dessa uppmärksamma användaren på information.

## Människa - maskinsystem

Ett människa-maskinsystem (MMS) är ett system där teknik och människor ingår och är i nära samspel med varandra. En del av de större komplexa tekniska systemen är också säkerhetskritiska, t.ex. avancerad sjukvård, kärnkraftverk och flyget. Detta innebär att misstag i dessa människa-maskinsystem kan vara livsavgörande.

I samtliga människa-maskinsystem är kommunikation mellan system och människa nödvändig. Ofta sker denna kommunikation via ett gränssnitt som består av dels displayer där operatören får information om systemets status, dels manöverdon för styrning av systemet. Ett exempel på ett vanligt människa-maskinsystem är bilen och dess förare. I gamla bilar var det lätt att lyssna efter avvikelser, felande cylindrar, högt eller lågt varvtal. I nya bilar är det nästan svårt att höra om motorn är på. Komforten ökar i en tystare miljö, bullret tröttrar

inte vid långkörning och det ges möjlighet till att prata med medpassagerare eller njuta av musik. Men man har också byggt bort information som man delvis idag, istället, försöker representera på en visuell display.

Även inom flyget är föraren isolerad från systemet och helt beroende av representerad information via gränssnitt. Man talar inom flyget att upp till 80 % av alla olyckor beror på mänskliga faktorn och många pekar då på piloterna. Men det som lett till ett felaktigt handlande kan ofta härledas till dålig utformning av gränssnitt och annan teknik som de använder.

Piloterna har en tydlig presentation av flygplanets status, position, och färdriktning och vid överskridna värden, avvikande händelser eller icke-normala situationer används ett speciellt varningssystem för att uppmärksamma piloterna på kritisk information. Det visuella gränssnittet står för mer än 70% med lampor, text, instrument och symboler. Audiellt representeras information av pip, signaler och röstmeddelanden.

Den mest uppenbara fördelen med att använda varningssystem är att systemen är så komplexa och högautomatiserade att det skulle vara en omöjlighet för en människa att kontinuerligt övervaka och kontrollera samtliga processer i ett system. Dessvärre är att operatören inte alltid är insatt när något inträffar. För även om varningen är tydlig kan det vara svårt att sätta sig in i ett sammanhang man inte följt över tiden.

### **Varningsljud**

Generella fördelar med att använda varningsljud är som nämnts tidigare att man inte kan bortse från ljud och att man hör ljud även om man inte direkt lyssnar efter det. Alla i cockpit får på så sätt samma information oavsett vad de gör och missar inte viktig information. Visuell information ser vi bara om vi tittar på den eller åtminstone har det inom synhåll. Nackdelar med ljud är andra sidan av samma mynt. Varningsljuden kan vara distraherande och stör kommunikation i övrigt.

I internationella standards står t.ex att varningar ska uppmärksamma operatören och möjliggöra en bedömning av allvarlighetsgraden, identifiera problemet och

lämplig åtgärd. I varningssystem är det ofta ljud som uppmärksammar piloten på att något skett och riktar pilotens uppmärksamhet till den visuella displayen där det finns mer detaljerad information.

Hur kritisk en varning eller ett larm låter, anses vara en egenskap hos ljud. Man har börjat identifiera hur vi bedömer hur kritiskt ett ljud låter och på så sätt förstå hur vi ska kunna utforma varningsljud med relevant grad av allvar. Intensitet är det som i högst grad förmedlar fara. Nu kan inte denna parameter användas i så stor utsträckning vid utformning av ergonomiska varningsljud eftersom risken att öka ljudets intensitet tills det blir störande och obehagligt är stor. Ett ljuds repetitionsfrekvens och oförutsägbarhet ökar också känslan av allvar i varningen. Detta kan användas för att man ska kunna skilja två lika kritiska varningar åt t.ex. med hjälp av olika rytm.

### **Exempel på användning av varningsljud**

Varningssystem är nödvändiga och används i många tekniska system. Här ges exempel från tre system som visar vanliga problem och hur varningsljuden kan fungera i praktiken.

### **Sjukvård**

Inom sjukvården är det idag vanligt med olika övervakningsteknik innehållande varningsljud. Varningssystemen indikerar status på patienten och på själva utrustningen. Vanliga parametrar som övervakas hos patienten är puls, syresättning och blodtryck. Varningsljuden är en blandning av elektroniska pip och ljud. Inom sjukvården måste man ta hänsyn till att inte störa patienten, om denna är vaken, eller att störa andra patienter i närheten.

Personalen menar att man i kritiska fall jobbar så nära patienten att man redan förutsett varningen innan den kommer. Man ser och lyssnar direkt på patienten.

Ett exempel på ett informativt omgivningsljud är det ljud som hörs när trycket i syretillförseln till respirator faller, man hör att "luften går ur" systemet.

Vid en genomgång av olika utrustningar på avdelningen syntes avnötta larm-avstängningsknappar och vid flera tillfällen under analysen noterade personalen att det fanns olika ljud, vilket de inte hade hört tidigare.

Ett riktigt dåligt exempel är en värmemadrass som nedsövda patienter ligger på för att hålla kroppstemperaturen. När madrassens larm går igång låter det värre än hjärtstilleståndslarmet och den larmar endast för att den är över 38 grader varm!

Mycket utrustning larmar ofta fel då patienten är mindre än normalt eller är ett barn. Det går att ställa in individuella värden men detta upplevs som långt ifrån enkelt och för tidskrävande.

### **Kärnkraftverk**

I en kärnkraftssimulator analyserades varningssignalerna som används i olika system. I ett kärnkraftverk larmar det så fort ett värde passerar det normala eller en förändring skett. Detta behöver inte nödvändigtvis vara oväntat eller onormalt.

Kontrollrummet är stort och olika systemdisplayer står utspridda i rummet. Varningsljuden är kraftiga och man kan i många fall lokalisera vilken panel som larmar. Är det flera larm blir det svårare och ljudnivån är hög.

Varningsljuden har två nivåer; en lägre som kallas "pling-plong" och en mer kritisk som kallas "kriget kommer", som ljuder om inte den första varningen hanteras inom 10 sekunder. Det senare larmet var som namnet antyder mycket obehagligt.

Vid revision av systemet passeras så många värden att man sätter en person på att bara återställa larm en hel dag!

### **Flygplan**

Ett flygplan har ofta ett varningssystem som installeras av tillverkaren. Detta har vanligtvis två nivåer varningsljud ett mindre kritiskt larm och en mer kritisk varning. Dessa uppmärksammar piloten på de visuella displayerna. Det finns

genom molnen 15 sek före nedslag visste piloterna redan att det bara var glidflygning som återstod. De meddelade flygtrafikledningen att de skulle kraschlanda. Trots denna situation så fortsatte GPWS "-Whoop whoop pull-up!" om och om igen och fortsatte ända tills de slog i backen...

I intervjuer har piloter givit dessa synpunkter kring varningsljuden i cockpit:

- Så många signaler är förvirrande.
- Ljuden går inte att skilja åt.
- Pling-plong, meningslöst!
- Klocka och horn är svåra att särskilja.
- Det skulle vara bra att kunna stänga av ljuden så man kunde tänka.
- Jag kan ibland inte höra radiotrafiken för alla ljud.
- Stäng av ljuden! Eller korta åtminstone av signalerna.
- Brandlarmet har alldeles för högt ljud.
- Mycket distraherande ljud.

I en varningssituation krävs alla de mentala resurser som finns att tillgå hos piloterna för att lösa det uppkomna problemet. Att irritera sig på ljud eller bli störd är onödig informationsbehandling.

### **Informationsbehandlingsprocessen**

Förenklat kan man säga att vi har en begränsad resurs som ska räcka till att uppmärksamma och uppfatta information, tolka och förstå informationen och kanske jämföra detta med tidigare erfarenheter som finns i minnet. När vi får ont om resurser kan vi inte överväga alternativa lösningar, får svårt att koncentrera oss och blir stressade. Ju lättare vi kan uppfatta information och ju mer vi känner igen eller kan associera till något, desto mindre resurser tas i anspråk.

I internationella rekommendationer och standards anges ofta att man ska utforma varningssystem och ljud som minimerar belastningen på informationsprocessen och minnesfunktionerna. Informationsbehandlingsprocessen inkluderar både att perceptuellt uppfatta varningsljudet och att kognitivt förstå dess betydelse och sammanhanget i vilken den uppkommer. När varningsljuden

om resurser kan vi inte överväga alternativa lösningar, får svårt att koncentrera oss och blir stressade. Ju lättare vi kan uppfatta information och ju mer vi känner igen eller kan associera till något, desto mindre resurser tas i anspråk.

I internationella rekommendationer och standards anges ofta att man ska utforma varningssystem och ljud som minimerar belastningen på informationsprocessen och minnesfunktionerna. Informationsbehandlingsprocessen inkluderar både att perceptuellt uppfatta varningsljudet och att kognitivt förstå dess betydelse och sammanhanget i vilken den uppkommer. När varningsljuden upplevs meningslösa, svåra att särskilja från varandra, och distraherande, så är detta ett resursslöseri, som i situationer som kräver problemlösning och kreativitet därmed kan äventyra säkerheten.

Problemen som identifierades med de varningssystem och ljud som används idag kan sägas höra till två kategorier. Dels vilken typ av ljud som används (Är de meningsfulla för operatören?) och dels hur dessa presenteras (Repeteras de för ofta och låter för länge? Går de stänga av eller är de för högljudda?). Många av operatörernas synpunkter vittnar om att det ofta är en kombination av dessa aspekter som bör förbättras.

De ljud som används idag är nästan uteslutande konventionella pip och signaler. Delvis beroende på att tekniken har haft begränsningar att skapa och använda ljud i systemen, delvis på grund av just konventioner. Varningsljud brukar låta så.

### **Naturliga ljud**

En pilot uttryckte sitt missnöje med "pling-plong meningslöst". I naturen finns bara ljudspektrum, inga rena toner. Innan stämgaflöj och oscillator uppfanns är det tveksamt om ett öra någonsin blivit stimulerad av en enskild frekvens eller hört en ren ton. Örat är troligtvis utvecklat för att fånga upp just inbördes förhållanden mellan olika frekvenser och mönster av ljud.

På liknande sätt kan man säga att minnet är uppbyggt för att associera till mönster och bilder av flera dimensioner snarare än att minnas och särskilja

endimensionella stimuli. Om många signaler är lika kan vi inte hålla isär mer än några stycken och när vi närmar oss gränsen till vår förmåga så blandar vi ihop signaler och blir förvirrade.

Utifrån detta resonemang gjordes experiment för att visa på det orimliga att använda signaler som påminner om varandra och är meningslösa som varningsljud istället för att använda mer komplexa och kanske meningsfulla ljud som kräver färre mentala resurser. Detta skulle i en krissituation frigöra mentala resurser för t.ex problemlösning och bidra till ökad säkerhet i människa-maskinsystem.

### **Associerbarhet**

I ett experiment jämfördes inläring av varningsfunktioner som tilldelats olika kategorier av ljud, alltifrån djurljud och ljudikoner till konventionella varningssignaler.

Försökspersonerna fick memorera varningsfunktioner som var tilldelade olika ljud i en kategori och memorera dessa. Efter det fick de para ihop samma funktioner med en annan uppsättning ljud från en annan kategori och memorera dessa. Experimentet gick till så att man fick en genomgång av funktion-ljudpar och sen spelades ljuden upp ett efter ett och man skulle då svara på vilken funktion de representerade. Om man fick fel fick man en ny genomgång osv. tills man kunde samtliga funktioner-ljudpar. Antalet genomgångar som krävdes var mått på hur lätta ljuden hade varit att associera till given funktion.

Resultaten visade att förutom kontrollgruppen där röstvarningar användes istället för ljud, så var det lättast att lära sig ljudikoner eller "naturliga" ljud som föreställde funktionen. Efter dessa var djuren lättast att associera till funktionerna. Djurens läten hade ingenting med funktionerna att göra men man kände igen ljudet och kunde på så sätt skapa en association eller bild av hur djuret och funktionen hörde ihop. Sist och svårast var det att lära sig de konventionella varningsljuden.



Lite provocerande kan man konstatera att ur denna aspekt skulle man istället för pip, piiiip och pip-pip med fördel kunna använda gnägg, grymt och vov.

Att använda naturliga ljud skapar mycket utrymme för val av ljud. Man kan inkludera dynamik av förlopp, höra om något ökar eller minskar, eller tagit slut. Naturliga ljud kan förmedla mycket information i sig, men det är viktigt att säkerställa att det i möjligaste mån är relevant information för användaren. Det är därför en begränsning att som ingenjör sitta och hitta på ljudikoner till funktioner i ett system då det finns operatörer som är experter på systemet och som vet vilka ljud som är naturliga inom systemet eller för användaren. Därför gjordes även följande experiment.

### **Ljudbildning**

Om man behöver utveckla ljud till ett antal varningsfunktioner kan man använda sig av metoden ljudbildning. Några funktioner i taget presenteras i skrift samtidigt som det presenteras en mängd ljud. Systemanvändarna får för varje funktion ange vilket ljud han eller hon associerar till denna funktion och även på en skala mellan 1 och 10 ange hur väl ljudet överensstämmer med deras "bild" av funktionen.

Denna metod har använts på helikopterpiloter. Det var intressant att se att de faktiskt var överens om flera ljudassociationer. I experimentet inkluderades även de ljud som finns i de helikoptrar de flyger. De flesta av de konventionella ljuden blev ratade av piloterna. Det är dock viktigt att notera att vissa av de nu existerande ljuden ändå föredrogs av piloterna även om de är klassiska pip och inte komplexa och i teorin meningsfulla. I praktiken var dessa ljud "naturliga" för piloterna i den miljö som de vistas i.

Om man med ljudbildningstest kan skapa en uppsättning varningsljud så kräver de ingen inlärning, man hör vad de innebär.

### **Presentationsformat**

Efter att ha nått lite längre i att förstå vilka typer av ljud som skulle kunna vara bättre än de konventionella pipen återstod frågan om hur dessa ska presenteras.

Flera av synpunkterna från användare antyder att ljuden spelas för länge eller upprepas för ofta och inte alltid går att stänga av.

I en studie gjordes en inledande simulering för att diskutera olika alternativ av presentations-format. Även här deltog helikopterpiloter och de hade ett uppdrag där de skulle hämta utsatta personer bakom fiendes linje vilket innebar att de flög lågt och snabbt för att undvika att bli upptäckta. Under flygningen fick de varning för att bränslet tar slut om 20 min och/eller att motorn havererar.

Piloterna fick flyga med följande presentationsformat: 1 varningspresentation, 2 repetitioner, 3 repetitioner, varningen repeterades tills den stängdes av manuellt och sist varningen repeteras tills felet är åtgärdat. Efter varje flygning fick piloterna svara på frågor och vikta olika förhållanden under testet.

Resultatet visar att det ofta beror på situationen när en varning blir mer störande än den ger relevant information. I en i övrigt lugn situation räcker det med en presentation. Är det något annat som pågår kan det vara bra med två eller fler repetitioner så man påminns inom kort tid, alternativt säkerställer att informationen gått fram. Piloterna föredrog att själva stänga av när man uppfattat varningen. I en lugn situation är detta inget problem, då stör inte en eller två varningar så mycket och man har tid att kvittera larmet men i en stressig situation kommer varje överflödiga repetition vara störande och samtidigt vill man inte slösa tid och resurser åt att stänga av ett larm. Ett förslag var att ha tre repetitioner åtföljt av en paus och sen upprepning av en, två eller tre repetitioner. Att ha kontroll och veta att det går att stänga av och hur länge larmet kommer att fortsätta minskar frustrationen och irritationen.

### **Avslutning**

Vid användandet av begreppet naturliga varningsljud inom människa-maskinsystem så ska man beakta det från två håll. Man bör använda ljud som är naturliga i den bemärkelsen att de är flerdimensionella och mer komplexa än enstaka pip och toner och därför mer lämpade för vårt hörselsystem. Men att endast använda ljud som återfinns i naturen är inte relevant i en teknisk miljö. Här finns andra ljud som är "naturliga" för användaren av systemet.

Vid konstruktion av ett varningssystem med ljud bör man ta hänsyn till bägge dessa aspekter och utnyttja den flora av ljud som skulle göra det lättare att särskilja larm från varandra, lättare associera till tilldelad funktion, lättare lära sig de olika ljuden och också göra det mer behagligt.

Mycket sällan har man nytta av ett varningsljud som får operatören att hoppa högt och sen inte heller kan tänka i en kakofoni av ljud. Varningsljud är ofta överdimensionerade i intensitet eftersom ingen eller lite hänsyn tagits till att anpassa frekvensintervallen för att undvika maskning, vilket innebär att ljud kan suddas ut andra ljud om de ligger i samma frekvensområde.

Visst finns det larm t.ex. överfallslarm och inbrottslarm där ljudet i sig ska vara avskräckande och väcka uppmärksamhet på långt håll. Men i många tekniska system kan man bära headset eller sitta nära och då är det i första hand information som behövs, inte distraherande stresshöjande ljud.

Oavsett ljudvolym så upplevs vissa ljud mer behagliga än andra. Kanske för att vi finner mening i vissa och inte i andra.

Denna forskning har fört teorin lite närmare praktiken och förhoppningsvis kan den leda till att vi i fortsättningen får höra mindre meningslösa pip i vår vardag eller på jobbet. Ljud kan tillföra något bra till många system, men det kan bli ännu bättre om de är av typen som gör lyssnandet mer roligt och meningsfullt.

## När ljud pockar på uppmärksamhet

Marcus Sanchez Svensson  
Blekinge Tekniska Högskola

### Introduktion

Ljud fungerar i många vardagliga sammanhang som en informationsbärare och ett praktiskt hjälpmedel. Vi är ständigt omgivna av ljud var vi än befinner oss: i hemmet, på bussen eller ute på staden. Det kan vara mikrovågsugnen som meddelar oss att maten är klar, bilen som informerar bilföraren om att bilbältet inte är fastspänt eller trafiksystemet som signalerar att det är fritt fram att gå över gatan på övergångsstället. Dock kan mängden förekommande ljud upplevas som om den pockar på vår uppmärksamhet och gör oss ibland undrande och sökande efter dess betydelse. Är det bara en upplysning eller är det en varning? Berör det mig och vad jag har gjort eller borde göra? Sådana praktiska tillämpningar ställer krav på ljudets utformning - in minst så att vi kan höra skillnad på olika ljud och att lära oss att känna igen dess specifika innebörd. Vi kan dock ställa oss frågan: Är det endast ljudets karaktär som gör att vi uppmärksammar ljudet och kan förstå dess betydelse och relevans? Denna artikel kommer att reflektera över hur ljud som händelse ges en innebörd för oss människor inte bara genom ljudets utformning utan också genom den sociala interaktionen, det praktiska sammanhanget och den fysiska omgivningen. Jag kommer att föra en diskussion som handlar om hur ljudsignaler som tillkallar och pockar på uppmärksamhet, i sig själv, är ett socialt fenomen – med andra ord, hur ljudvågorna får en social betydelse.

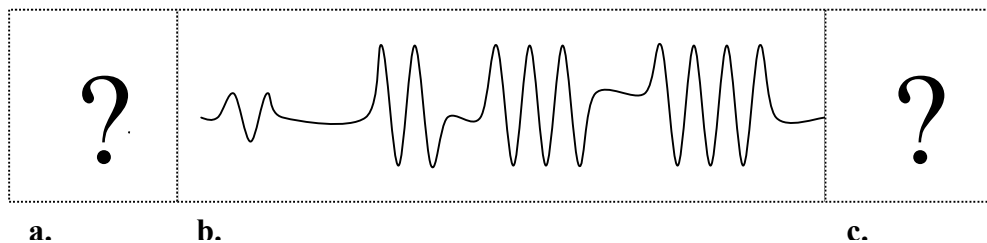
Diskussionen som jag för i denna artikel bygger på min forskningsbakgrund inom områdena arbetsvetenskap och sociologi. Under de senaste åren har jag bedrivit forskning som handlar om att studera människors vardagliga handlingar och aktiviteter i olika sammanhang och miljöer med ett specifikt intresse av att analysera hur människor använder verktyg och informationsteknologi. Min diskussion baseras på observationer och erfarenheter från en rad olika studier inriktad mot hur teknik stödjer arbete och samarbete i komplexa sociotekniska arbetsmiljöer. Dels handlar det om larmhantering på en dialysavdelning i Sverige och dels om erfarenheter från en arbetsplatsstudie av operationsrum på ett sjukhus i centrala London. I båda dessa studier har det uppdagats ett intresse av att fundera på hur människor i sociala och publika sammanhang hanterar händelser och situationer när ljud pockar på uppmärksamheten. Jag kommer att börja med en kort bakgrund till hur vi kan analysera och konceptualisera ljud som ett socialt fenomen. Jag

kommer sedan att ge några exempel från en av mina studier och diskutera hur ljud fungerar som informationsbärare.

De bör påpekas att det finns en del redan skrivit om ljud som informationsbärare och hur människor lär sig att lyssna på ljud (Gaver 1986; 1997, Alexanderson 2002) och om ljud i larmsystem och dess konsekvenser (Seagull och Sanderson, 1998; Momtahan, Tansley och Hetu, 1993; Meredith och Edworthy, 1994; Kesting, Miller och Lockhart, 1998).

### Ljud som socialt fenomen: tillkallare och svarare

Ljud som i olika former och sammanhang pockar på uppmärksamhet är ett intressant socialt fenomen. Frågan för mig är: Vad är det som egentligen *händer* när ljud pockar på uppmärksamhet, och *hur*? För att svara på den frågan så är det relevant att fundera på de sammanhängande händelserna innan (a, se bild nedan), under (b), och efter (c) ljudsignalen. Låt oss en stund fundera på en vardaglig händelse: en telefonsignal.



Emanuel Schegloff (1968) är en forskare inom sociologi som i en studie av telefonsamtal pekat på att ringande telefoner har en betydelsefull egenskap, nämligen att de är tillkallare, och att sådana kräver ett svar. Schegloff menar i sin analys att telefonsignalen och svaret bildar en händelsesekvens där en tillkallande telefonsignal lämpligen följs av ett svar - någon som lyfter på telefonen och svarar. Denna villkorliga händelseföljd är nog alla bekanta med som har hört en ringande telefon och känt uppmaningen att avbryta ett göromål för att kunna svara. Schegloffs studie lyfter också fram en annan fråga om hur bildandet av det sekventiella förhållandet mellan tillkallare och svarare kan identifieras och kännas igen. Schegloff observerar att personerna i närheten av en ringande telefon gör vissa tolkningar och analyser för att avgöra *vem* som ska svara, och *hur* och *när* (Maynard och Clayman, 1991). Låt oss ta ett bekant exempel.

Tänk er själv in i scenariot på arbetsplatsen när du och dina kollegor håller ett samtal i korridoren när det plötsligt hörs en ringsignal. Efter några ringsignaler vänder sig några om och tittar ner längs korridoren. Du vänder dig

om mot dina kollegor och ser att någon vänder sig om, höjer på ögonbrynen eller rycker på axlarna. Du tänker kanske efter en stund när inte signalen tystnar: "Vems telefon är det som ringer egentligen?", "Finns det inte någon som kan svara?". "Men gud! Det är kanske min telefon som ringer?". Precis när du ska rusa ser du att någon kommer springande och dyker in i ett av rummen i korridoren. Efter en stund tystnar signalen - tillkallaren har fått ett svar från en mottagare. Vi kan tänka oss många andra liknande scenarier så som till exempel när mobiltelefonen ringer på bussen eller på föreläsningen, med all den uppståndelse som det kan innebära. Poängen som jag vill göra här, och som genomlysas i Schegloffs studie, är att *ringsignalen blir ett socialt konstruerat objekt i och med att den i sitt sammanhang medför en tillkallare och en mottagares handling.*

En annan forskare inom sociologiämnet, vid namn Harold Garfinkel, har demonstrerat denna poängen genom en övning med sina studenter (se Garfinkel och Wieder, 1991). Garfinkel bad sina studenter att göra ljudinspelningar av ringande telefoner som är a) hörbart tillkallande bara dem själva b) hörbart tillkallande någon annan c) simulerat hörbart tillkallande bara dem själva d) simulerat hörbart tillkallande någon annan, och e) bara ringande. Det finns en lång utläggning om övningens resultat och hur studenterna reagerade på uppgiften, men jag skulle vilja nämna bara ett par utmärkande observationer. Först och främst ville Garfinkel demonstrera med denna övning att för att en telefon ska ringa så måste en aktör på ett eller annat sätt ge upphov till att en ljudsignal bryter tystnaden. Dessutom, att telefonen hörbart tillkallar dig, någon, eller ingen avgörs som ett resultat av deltagarnas aktiviteter vid det unika tillfället (Maynard and Clayman, 1991). Även om det bara är ännu en ringande telefon som hörs så gör själva lyssnandet efter signalen och de nuvarande omständigheterna – som till exempel att ringsignalen kommer ifrån en annan riktning - ändå att den "söker" efter någons svar. En annan aspekt av det som uppstår i händelsen av en påtagligt tillkallande telefon är att vi kan fundera på hur en deltagare tillsammans med andra kan interagera och påverka det sociala sammanhanget så att den "någon" som telefonsignalen tillkallar blir tydligare. Vi kan till exempel tänka oss i vårt tidigare scenario på arbetsplatsen att någon har sagt eller säger att "Chefen skulle snart ringa mig?" en stund innan telefonsignalen ljuder i korridoren. Denna sociala handling gör att telefonsignalen och handlingssekvensen tillkallare-svarare i just denna unika situation skapar ett sammanhang som tydliggör den möjliga mottagaren och svararen – snarare än att det handlar om en hörbar skillnad i termer av ljudets karaktär. Med andra

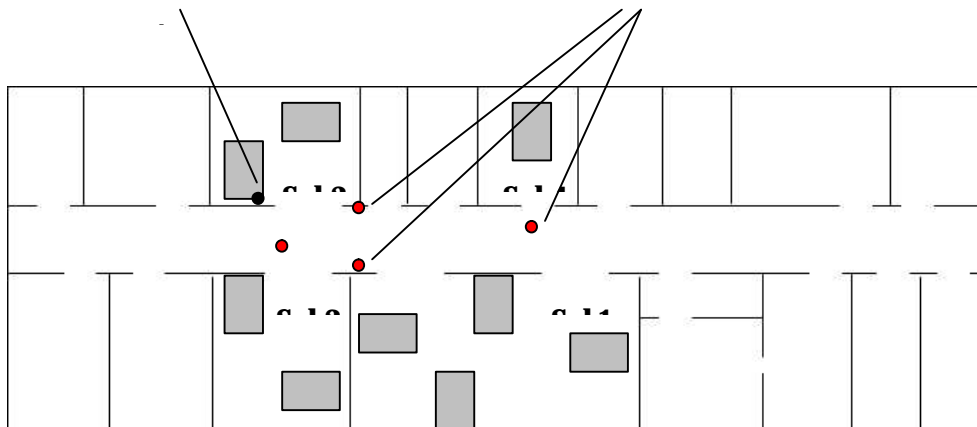
ord, vem signalen ljuder för skapas av de inblandade deltagarna vid det unika tillfället i ett socialt sammanhang.

Analysen av hur en telefonsignal pockar på uppmärksamhet och ges betydelse och mening i sitt sociala sammanhang tror jag är relevant i fundering på andra typer av vardagliga artefakter som på ett eller annat sätt bygger på ljudresurser. Ett sätt att fundera vidare på ljudets egenskaper är att studera hur människor som har det som en del av sitt yrke att reagera på ljudsignaler. Jag ska därför nu ge ett par exempel från mina tidigare studier där yrkesverksamma människor agerar på ljud som en del av deras vardagliga arbete – nämligen larmhantering i sjukhusmiljö.

### **Ljud blir larm: larmhantering på en dialysavdelning**

Jag skulle nu vilja beskriva ett par sekvenser från en observationsstudie på en arbetsplats som på många vis kan anknytas till diskussionen om hur ljud fungerar som en informationsbärare. Studiens syfte var att undersöka hur sköterskorna på en dialysavdelning hanterar dialysapparaternas larm (se Svensson och Tap 1998, Sanchez Svensson och Tap 2003). Det som var upprinnelsen till vår studie var en händelse som utspelades i sjukhuskorridoren på dialysavdelningen en eftermiddag. Låt mig först säga några ord om dialysavdelningen och larmanordningen.

På dialysavdelningen tar man emot patienter som tre dagar i veckan behöver genomgå en dialysbehandling – en behandling som innebär en rensning/filtrering av patientens blod som inte längre njurarna klarar av. Under behandling kopplas patienten till en dialysapparat som pumpar patients blod i ett kretslopp. I detta kretslopp finns ett dialysfilter som tjänstgör som en konstgjord njure. En förutsättning för att dialysbehandlingen både ska vara säker och reglerbar är att apparaten utför en mängd kontroller. När apparaten upptäcker en avvikelse i dialysbehandlingen signaleras detta genom ett larm. Apparaten avger då larmsignaler och aktiverar en huvudlarmlampan som börjar blinka. På skärmen beskrivs larmet antingen i ett larmfönster och/eller i ett av larmkällans monitorfönster. När sköterskan hör ett larm från maskinen kan han/hon utifrån skärmen ta beslut, åtgärda och kvittera larmet. Under själva behandlingen är patienten placerad i en av de många salarna utmed en lång korridor (se bild nedan).



Under ett normalt dialyspass tar fyra sköterskor hand om 8 patienter. Där av följer att under ett pass ansvarar varje sköterska för två patienter placerade i en sal.

Eftersom sköterskorna inte alltid kan vara närvarande i patientsalen så har avdelningen ett larmsystem i korridoren som gör larmsignalen hörbar från andra rum på avdelningen. Genom en larmsladd från dialysapparaten leds signalen ut till avdelningens larmlampor och högtalare. Utanför varje rum i korridoren finns en larmpanel antingen på väggen eller i taket utanför patientsalen (se bild ovan). Larmpanelen signalerar när maskinerna respektive patienterna inne i salen larmar.

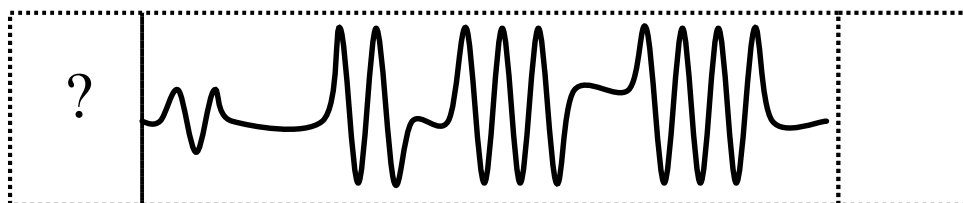
Under vår studie av dialysapparaternas larm utspelades den händelse som kom att ligga till grund för vår inriktning i studien. Det är strax innan middag före det andra arbetspasset går igång. Patienterna är inkopplade, vissa sovande, andra i samtal med varandra, och sköterskorna arbetar med sina sysslor. Plötsligt bryts lugnet av ett larmljud ute i korridoren. Det ljuder högt ute i korridoren och den röda larmlampan på väggen utanför den stora patientsalen blinkar. Efter ett par sekunder från det att larmet startade tittar en av sköterskorna ut från kontorsrummet. Ett par andra sköterskor tittar sig



också omkring stående utanför kontoret. Ännu en sköterska kliver ut i korridoren från lunchrummet och söker med sin blick i korridoren. Sköterskorna utanför kontoret vänder sig till varandra undrande: "Vem är det som larmar ... vem är ansvarig?", säger en av dem.

Lika plötsligt som larmet kommer är det borta. En av sköterskorna som var på väg mot salen tittar in i patientsalen innan hon återvände och visar ett leende mot någon inne i rummet. Under den korta stund som larmet signalerade tittade ett flertal sköterskor ut i korridoren. Larmljudet var högt och larmlamporna blinkade. Det var många frågor som dök upp hos oss under observationen av händelsen. Varför reagerade så många på larmet? Varför gick de inte direkt på larmet? Var fanns den ansvariga sköterskan? Det vi förstod senare var att detta var en vanlig vardaglig företeelse.

Ett larm kan vid första anblicken verka vara en ganska otvetydig varning eller tecken på att något är fel eller att det finns en omedelbar fara som någon måste reagera på. Larmet ingår i ett tekniskt system som varnar och uppmärksammar människorna i omgivningen. Men tittar man närmare på sådana situationer lägger man märke till hur människor lär sig lyssna på ljudet och skapar metoder/strategier för att hantera sitt eget och andras agerande i händelser av ljudlarm. Låt oss börja med en diskussion om skeenden och händelser innan tystnaden bryts (se bild nedan)

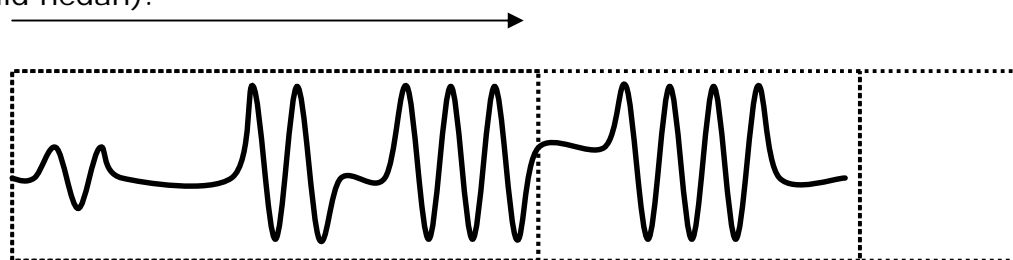


Frågeställning om hur ljudet pockar på uppmärksamhet har att göra med, som vi tidigare poängterat, att ljudsignalen i vissa sammanhang verkar som tillkallare och "utpekare" personalen som eventuella mottagare. Ungefär på samma sätt som telefon måste ha en aktiv tillkallare som ger upphov till en signal (Garfinkel och Wieder, 1991), så ljuder larmsignalen med specifika anledningar som kan härledas till det som omger dialysbehandlingen – en maskin som kopplas till en patient som behandlas och vårdas av en eller flera sköterskor. Ljudsignalerna aktiverats av maskinen (när det gäller maskinlarm) som upptäckt en avvikelse utifrån ett mätvärde som den är programmerad att signalera som en varning. Denna avvikelse har oundvikligen angivits i ett tidigare stadium av en mänsklig aktör som bestämt rent numeriskt avvikelens gränsvärden och att en sådan avvikelse måste larmas och kontrolleras. Det

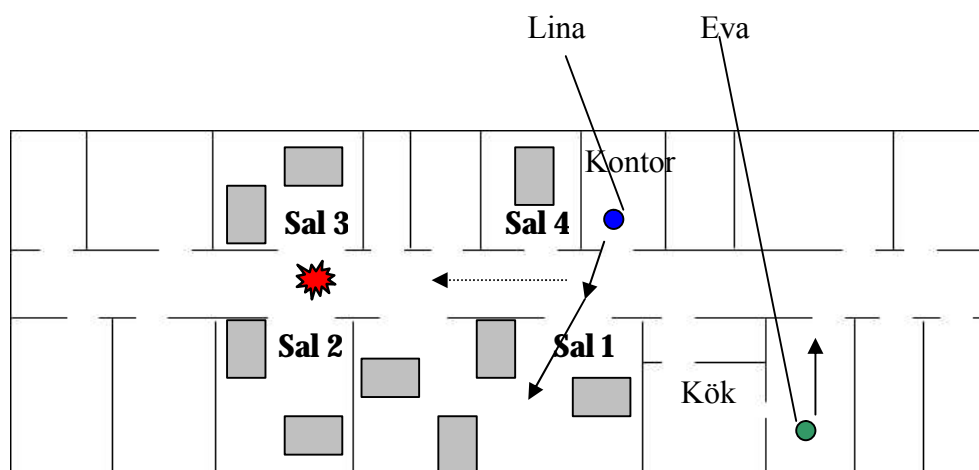
finns alltså en bakomliggande tillkallare som bryter tystnaden med en ljudsignal som personalen i miljön inte bara blir en mottagare av utan också en ansvarig svarare.

Det som är intressant att fundera på är hur personalens erfarenhet, deras koordinerade arbetsorganisation och det unika tillfället kan avgöra eller stå som bakgrund till förståelsen av vem som tillkallas och vem som borde agera vid det unika tillfället. På en del arbetsplatser, där ljudlarm är en del av vardagen, kan man upptäcka hur deras arbetspraktik har formats efter larmsystemets möjligheter och begränsningar. I observationerna på sjukhuset och i samtalen med personalen visade det sig till exempel att larmen kunde få olika betydelser beroende på de unika sammanhangen och att de själva agerade utifrån sådana uppfattningar. En intressant observation var att *tidpunkter och föregående händelser* påverkade sköterskornas lyssnande och agerande i larmsituationer. Vi uppmärksammade till exempel att larmsignalerna i korridoren under uppstarten av dialysbehandlingarna inte hade samma effekt som under pågående behandling. Anledningen var att alla då visste att maskinerna larmar eftersom behandlingen inte kommit igång, att de inte hunnit ställa in gränsvärdena, och att alla ändå är upptagna inne hos patienterna. Senare under dagen när en sköterska exempelvis går på lunch talar han eller hon om detta för någon av de andra sköterskorna så att de "täcker upp" för honom/henne. Då agerar den informerade sköterskan likadant på den frånvarande sköterskans larm som han/hon gör på sina egna. Många gånger kunde vi också i våra observationer höra hur sköterskorna i lunchrummet eller på kontoret berättade för varandra om problemen som de hade eller kunde förvänta sig med en viss patient. De kunde kanske säga: "Idag har jag haft problem med Lennarts blodflöde igen. Den maskinen kommer att larma hela dagen". Sådana diskurser och kontinuerliga samtal mellan sköterskorna stod som en bakgrund till sköterskornas agerande i larmsituationer. Sköterskorna lyssnade och reagerade på ljudsignalerna baserat på en vetskap och erfarenhet om maskinernas egenskaper, patienternas tillstånd, deras egna insatser och den koordinerade arbetsfördelning, just där och då, vid de unika tillfällena.

En annan aspekt av larmhanteringen som observerades var att *det tar tid för ett larm att bli ett larm* – alltså skeendet under tiden som signalen ljuder (se bild nedan).



Vad menar jag då med att det tar tid? Föreställ dig att det hörs ett larmljud i korridoren. Som sköterska vet du att det kan vara ett falskt larm eller ett övergående larm som du inte behöver bry dig om just nu eftersom t.ex. ett gränsvärde kanske temporärt överskridits. Du lyssnar därför och hör om ljudet fortsätter. Det kan ju också vara så att den sköterska som är ansvarig för patienten som det larmar hos redan finns i patientsalen (som du annars skulle behöva assistera). Sådana faktiska vanliga företeelser gör att sköterskorna inte alltid rusar mot varje larm direkt. Ett endaste larmljud blir inte alltid ett larm för sköterskorna. En upprepning i en viss sekventiell följd är snarare det som i vissa sammanhang har betydelsen av ett larm – betydelsen tar tid. Låt oss ge ett exempel.

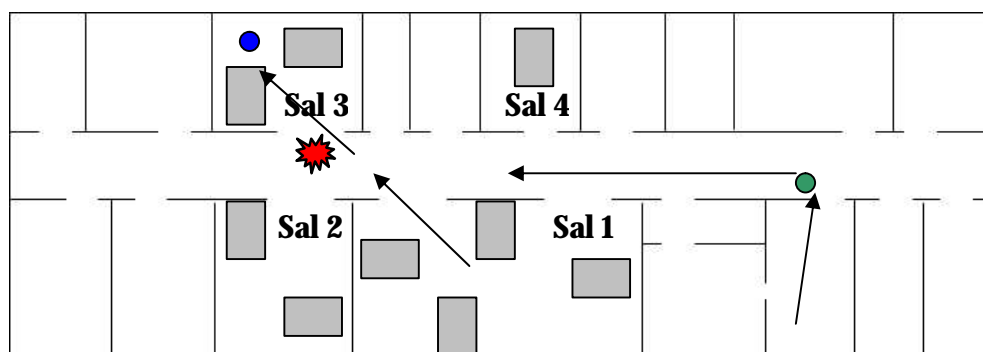


Figur 1: Utgångspunkt för Linas och Evas agerande

Under ett dialyspass hade sköterskan Eva två patienter i sal 3, medan den andra sköterskan Lina hade två patienter i sal 1. Vår observation börjar när Eva står i lunchrummet (se figur 1) och pratar med en kollega. Lina befinner sig vid detta tillfälle på kontoret. Både Linas och Evas patienter ligger uppkopplade till

dialysapparaten sedan några timmar tillbaka. Vid denna tidpunkt går ett larm från en maskin i sal 3, vilket som vanligt också signaleras ute i korridoren.

Larmet går samtidigt som Lina i lugn takt kommer ut från kontorsrummet och går över korridoren in i sal 1 (se figur 1). När hon befinner sig mitt i korridoren tittar hon mot larmlamporna och ser att det är larm i sal 3. Efter ytterligare 4-5 sekunder kommer hon ut ur sal 1 och går, fortfarande lugnt, mot sal 3 där det larmar. Under tiden har också Eva uppmärksammat larmet och börjat gå mot salen. Lina får efter en stund tyst på larmet inne i salen samtidigt som Eva kommer halvspringande i korridoren (se figur 2). Fastän hon ser att Lina redan är i salen går hon också in.



Figur 2: Lina tystar larmet medan Eva tar sig mot salen.

Sköterskorna agerar olika på larm beroende på om det är i deras egen sal eller inte. I videoklippen ser man att Lina inte reagerade direkt på larmet eftersom det inte är i hennes sal. För Lina blev det inte ett "riktigt" larm förrän efter cirka 4-5 sekunder. Anledningen till att hon väntade berodde antagligen på att hon ville veta ifall den ansvariga sköterskan var i närheten och kunde ta hand om larmet själv. Om sköterskan upptäcker att det är i hennes sal det går ett larm, reagerar hon alltså fortare än annars (såvida hon inte vet att den ansvariga sköterskan är upptagen). De har också själva, i intervjuer efteråt, förtydligat våra observationer med att säga att de faktiskt avvaktar innan de går på ett larm eftersom det förekommer "falska" larm. Med ett falskt larm menar vi ett larm som automatiskt återställer sig själv efter ett par sekunder. Ett exempel på detta är när dialyskoncentrationen i apparaten tillfälligt går över eller under önskvärd nivå för att sedan leta sig tillbaka till normala värden igen.

En annan dimension av att höra ljud och förstå dess innebörd är faktiskt att kunna *se och observera* något. Som ni kanske har märkt så har jag i texten presenterat sådana fall och exempel där det är svårt att skilja på ljudupplevelsen och vad som kan uppfattas i den omedelbara fysiskt synliga omgivningen. Att höra vem som ljudet signalerar till eller vad dess innebörd är

sker oftast i kombination med synliga resurser. Det handlar då inte bara om att ljud ofta kombineras med visuella presentationer (som blinkande lampor) utan också om hur vi människor använder varandras agerande gentemot varandra och den fysiska omgivningen (inkluderat ljudet) som resurser. I våra föregående exempel har vi sett hur människorna har inte bara lyssnat sig till en förståelse utan använt observationen av hur andra reagerar på ljudhändelser och vad som händer vid sådana tillfällen. Till exempel så har vi sett hur sköterskorna springer ut i korridoren inte bara för att se larmlamporna men också hur andra sköterskor agerar och visar att de själva uppmärksammat ljudet/larmet och varandra.

Till detta hör naturligtvis också, vilket inte ska förringas, att personalen kunde lära sig att höra det som är mer hörbart. Exempelvis sade de sig använda möjligheten att höra varifrån ljudet kommer. Ibland påpekade de att de agerade direkt då de kunde höra att det var deras patienter eftersom det larmade längre ner i korridoren. De observerade också att larmsystemet hade vissa egenheter när larmet stängdes av. Ibland förvrängdes ljudet när någon aktivt stängde av larmet på maskinen tillskillnad från när larmet stängdes av själv. Sådana rena ljudegenskaper gav personalen ytterligare resurser i deras larmhantering och dialysarbete.

### **Diskussion och summering**

Det jag har försökt lyfta fram till diskussion och fundering är vad som egentligen är bärare av information. Är det ljudet i sig själv eller helt enkelt resultatet av interaktionen mellan oss människor? Jag har velat poängtera att ljud som pockar på uppmärksamhet, och som fungerar som informationsbärare, är ett gemensamt socialt åstadkommande. Det är individen tillsammans med alla andra som påverkar den sociala miljön som sedan står som bakgrund till hur vi lyssnar och agerar när ljud pockar på uppmärksamhet. Det kan då vara intressant att fundera på hur vi kan designa ljud inte bara med den enskilde lyssnaren som modell utan även med individen i ett socialt, praktiskt och fysiskt sammanhang.

Några av de sakerna som man skulle kunna peka på i ett design sammanhang är att en djupare förståelse av det sociala kontext borde vara en integrerad del av ljuddesign – i alla fall i den faktiska tillämpningen. I vårt exempel har vi sett hur ljudet pockade på uppmärksamhet och ledde till vissa agerande influerat av det faktiska dialysarbetet och hur de lärt sig att koordinera sina arbetsuppgifter. Det står då nära till hands att fundera på hur tillämpningen av ljud kan gå hand i hand med det arbete som tillämpningen ska

stödja och den arbetsorganisation som växer fram. En konsekvens av detta skulle kunna vara att man inser att vissa ljudegenskaper eller sätt att implementera ljudresurser skulle kunna stödja istället för att hindra framväxten av ett smidigt koordinationsarbete.

Frågeställningarna i denna artikel är speciellt relevanta för mig i ett vidare informationsteknologiskt framtidsperspektiv. I vår vardag omges vi av datorkraft som inte längre bara sitter bekvämt på våra skrivbord utan finns nu mer och mer tillgänglig var vi än befinner oss: t.ex. i den handhållna enheten, i bilen, i vardagsrummet, eller i affären. Inom forskningsvärlden pratas det mycket om en vision som kallas 'Ubiquitous Computing' (Weiser, 1994, 2002) – allestädes närvarande datorkraft – och hur vi kan se framför oss en värld av datorer som gör sig tillgängliga var vi än befinner oss på ett enkelt och smidigt sätt. Det pratas om hur information och datorkraft skulle kunna göras tillgänglig i t.ex. väggen, möblerna, kläderna eller kaffekoppen. En av de stora utmaningarna med en sådan vision är hur att fundera på hur vi kan undvika ett inferno av datorer som ständigt pockar på vår uppmärksamhet med signaler och information om det ena och det andra? Ljud kan säkerligen komma att bli en viktig resurs för att underlätta relationen mellan datorn och användaren – hur detta kan ske är dock en frågeställning som måste utforskas.

## Referenser

Alexandersson, P. (2004). Peripheral awareness and smooth notification: the use of natural sounds in process control work. Publicerad i den tredje Nordiska Human-computer Interaction konferensen, NordiCHI'04.

Garfinkel, H. och Wieder, D.L. (1991). Evidence for locally produced, naturally accountable phenomena of order, logic, reason, meaning, method, etc., in and as of the essentially unavoidable and irremediable haecceity of immortal ordinary society, IV: Two incommensurable, asymmetrically alternate technologies of social analysis. I Watson, G. Och Seiler, R. Text in context. Beverly Hills: Sage.

Gaver, W.W. (1997). Auditory interfaces. In Helander, M.G., Landauer, T.K. & Prabhu, P. (Eds.), *Handbook of Human-Computer Interaction*, 2nd edition. Amsterdam, The Netherlands: Elsevier Science.

- Gaver, W.W. (1986). Auditory icons: Using sound in computer interfaces. *Human-Computer Interaction*, 2, 167 - 177.
- Kesting, I.G., Miller, B.T. & Lockhart, C.H. (1998). Auditory alarms during anaesthesia monitoring. In *Anesthesiology*, 69, 106.
- Meredith, C. & Edworthy, J. (1994). Sources of confusion in intensive therapy unit alarms. In N.A. Stanton (ed). *Human factors and Alarm design*, pp. 207-219.
- Momtahan, KL, Tansley, BW. & Hetu,R. (1993). Audibility and identification of auditory alarms in the operating room and intensive care unit. In *Ergonomics*, Vol. 36, pp. 1159-76.
- Sanchez Svensson, M. (2005) *Configuring awareness: work, interaction and collaboration in operating theatres*. Ph.D. thesis. University of London.
- Sanchez Svensson, M. and H. Tap (2003) *Alarms – Localisation, Orientation and Recognition*. *International Journal of Human-Computer Interaction*. 15(1), 51-66. Laurence Erlbaum Associates, Inc.
- Svensson, M. och Tap. H. (1998). *Dialysavdelning 28: Utformningen av ett larmsystem*. Magisterarbete på Blekinge Tekniska Högskola (BTH).
- Schegloff, E. A. (1968). Sequencing in conversational openings. *Am. Anthropology*, 70: 1075-1095.
- Schegloff, E. A. (1986). The routine as achievement. *Human Studies* 9: 111-51.
- Seagull, F.J. & Sanderson, P. (1998). Anaesthesia alarms in surgical context. In *Proceedings of the 42th annual conference of Human Factors and Ergonomics Society*, Chicago, October 4-7.
- Weiser, M. (2002) *The Computer for the 21st century, Pervasive Computing*, Vol.1, No. 1,(January –March 2002): 18-25.
- Weiser, M. (1994). The world is not a desktop. *Interactions*, January.

## **Playing with auditory interfaces**

Dan Gärdenfors 2006

### **Introduction**

In most computer games, information is mainly conveyed by graphics while sounds generally are added as embellishment. However, in sound-based computer games, often referred to as “audio games,” all game contents are communicated sonically, including all objects inside the game world, instructions and menu options. This type of games has so far been developed mainly for blind and visual impaired (BVI) users. (Extensive information on numerous BVI-accessible games can be found at a website specialising in audio games [1]). Like most other computer applications developed for visually impaired users, audio games often involve a lot of recorded or synthesised speech. However, as speech generally is not an adequate substitute to graphics, there are many audio games that instead attempt to use of “illustrative” sounds that correspond more closely to the visual contents of modern computer games. Research within the Human-Computer Interaction (HCI) field has shown that non-speech auditory interfaces can be useful in many different types of computer applications [2]. This text is based on my experiences from two projects that dealt with audio game development. By incorporating ideas from visual design and contemporary musicology, I aim to present a wider perspective on sound design that can be useful to developers of various types of applications.

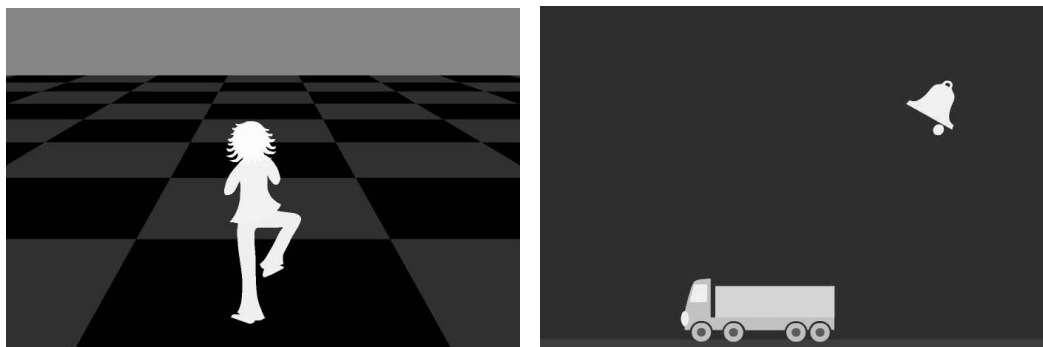
### **Audio games**

How can people with visual impairment play computer games that offer similar experiences to those of graphic games? Most digital media for the BVI community is text-based and communicated via speech synthesis or speech recordings. When designing computer games based on large chunks of recorded or synthesized speech there is a risk that the outcome more resembles old text-based games than modern computer games with animated graphics. So, is it possible to create games where sound effects and music replace graphics? This question was targeted in two game design projects. The first one was organized by TPB (Talboks- och Punktskriftsbiblioteket, the Swedish Library of Talking Books and Braille) and resulted in four, small, web-based games [3]. The second was an EU funded project called TIM (Tactile Interactive Multimedia), which



produced several audio game prototypes and one completed game called *Mudsplat* [4]. All games can be downloaded or played online, so they will only be described briefly in this text.

The TPB games are small Macromedia Flash™ applications, aimed at BVI users who have little previous experience of playing computer games. The games are published on the TPB web site to ensure that they can be accessed on various computer platforms, as long as there is an Internet connection and a Flash-enabled web browser. The games belong to two categories: the action games *Tag* and *Skybells*, and the puzzle games *Towers of Hanoi* and *Memory*. *Tag* is a game where the player chases a character that tries to escape by turning right and left in a random manner. In *Skybells*, the player controls a truck that catches bells that fall from the sky while avoiding falling rocks. *Towers of Hanoi* and *Memory* are adaptations of existing puzzle games. In addition to auditory interfaces, all these games also feature a type of high contrast graphics used in other TPB productions (see Figure 1). This means that blind children can play together with friends with partial or full sight, who can choose to use the graphics if they find the sound-based gameplay too difficult. (See [5] [6] for extended analyses of the TPB games)



**Figure 1.** Images from *Tag* and *Skybells*. Graphics by Annica Norberg, courtesy of TPB [3].

The TIM game *Mudsplat* is more complex and uses a lot of the conventions from the arcade game category, with features such as “bosses”, “extra lives” and “power-ups”. In *Mudsplat*, the player controls an avatar that defends the Earth from mud-throwing monsters. The player can reach 25 different game levels by defeating monsters and the difficulty increases throughout the game. The player moves the avatar from left to right in a virtual auditory space, trying to locate,

aim and shoot at monsters before they throw mud. The sounds of the monsters are partially concealed by a background music loop, which increases in intensity and complexity as one progresses through the game. (See [7] for an extended analysis of the TIM games and prototypes)

### **Sound sources**

When developing applications based on auditory interfaces, there are several kinds of sounds that can be used. Put simply, the main categories used are speech, "sound effects" and music.

Recorded or synthesised speech is useful for conveying very precise information, such as game instructions. However, within a game, speech tends to grow tedious if repeated too often. It is also generally too slow to communicate events occurring at the high tempo that characterises many types of action games. Neither is speech suitable if more than one message is conveyed simultaneously. This means that in many audio game contexts, it is often better to illustrate objects and events with non-speech sounds.

Within HCI, there are traditionally two main approaches to developing non-speech auditory interfaces. One alternative is to use sounds that are somewhat realistic, or iconic, such as using the sound running water to illustrate a water hose in *Mudsplat*. The other option relies on more arbitrary mapping by using short musical phrases to represent different objects and events, e.g. the guitar chord that indicates that one has succeeded in *Tag*.

The first, iconic approach, as introduced by Gaver (e.g. [8]), has many advantages. Gaver has shown that "auditory icons" in general are easy to learn and to recognise since they build on users' previous experiences. However, it is often difficult to find authentic sound recordings that can be used in interfaces, since many objects and events do not generate sounds in the real world. Gaver's solution to this is to use sounds that are slightly related to what they are to represent and modify them to fit into the auditory context. This design approach is used extensively in the TPB games, e.g. in *Towers of Hanoi*, where stone discs make "rocky" sounds, and the wooden poles sound "woody." This design method is closely related to visual iconography. If an image can say more than a

thousand words, visual iconography is about reducing attributes from an image, until it only says one or a few words, such as “camping” or “restaurant”. The cartoonist Scott McCloud describes the expressive iconic language of cartoons as “amplification through simplification” [9]. I believe that, when done well, this also applies to sounds and that iconic sounds in a sense can be much more expressive than authentic recordings. For example, to me there is more bounce in an artificial “boing” than in most authentic sounds of bouncing objects. This is why several sounds in the Mudsplat and Tag games are designed to resemble cartoon music and sound effects, similar to those by Carl Stalling and Treg Brown at Warner Brothers [10].

The second approach to auditory interface design is to create more abstract auditory messages from small “musical” phrases, often using instrumental sounds. This type of auditory symbols, often called “earcons” [11], show little or no resemblance to what they represent. This means that earcons can be used to indicate any phenomenon imaginable. Earcons can also be designed to generate pleasant, “musical” interfaces, since sounds can be chosen for their aesthetic qualities and how well they can be combined. The main drawback of earcons is that they can take a long time to learn, as they are more arbitrarily linked to the real world than auditory icons are. However, by incorporating musical conventions, such as using major harmonies to indicate success, earcons can make some use of listeners’ previous experiences. In the TPB Memory game, simple pitched percussion sounds are used to represent checking positions in a grid, since this event is not clearly associated with a sound in the real world. Other events in Memory that are musically illustrated include the rewarding sounds played when finding a pair of cards or winning the game, and the error alert sound heard when attempting an impossible move.

### **Positioning and timing of sound objects**

When creating BVI-accessible computer games, it is very difficult to adapt existing computer games by simply adding more sounds. Instead, to develop a comprehensive audio game it is better to plan the entire game concept around events and objects that can be represented sonically. Since the way we perceive sounds is very unlike the way we acquire visual impressions, both with regard to

space and time, audio game development requires a fundamentally different design approach to that of graphical games.

Most home computer systems only include stereo speakers or headphones. This means that there is only one spatial dimension, left to right, that can be used for positioning sounds, compared to the two dimensions of a computer screen. Both Mudsplat and the TPB games rely heavily on stereo positioning to separate sound objects spatially. In Mudsplat and Tag, the soundscape is presented from a first person perspective, as if the player was in the same place as the avatar. Memory and Towers of Hanoi are based on a third-person perspective. With only the left-to-right dimension available, other spatial dimensions must be simulated, either by a "realistic" decrease in sound pressure as objects are further away from the avatar in Mudsplat, or by including musical conventions such as the mapping of pitch to height in Skybells and Memory. Much more realism could be achieved with HRTF (Head Related Transfer Function) [12], although this has not been included in any of these games.

A key issue when developing auditory interfaces is to give the listener an overview of all objects present in the virtual environments. In everyday life, not all objects emit sounds and most sounds have limited durations, which can make it difficult for a blind person to overview present objects. When creating an audio game, this issue can be solved, although this generally results in "non-realistic" soundscapes. One option is to make all objects in the game space emit continuous sounds so the player can hear all objects within hearing distance at any given moment. This type of "continuous display" solution has been used in a Towers of Hanoi application by Winberg and Hellström [13]. While this approach provides the player with a constant overview, the simultaneous sounds risk blending into a complex and often noisy soundscape.

The TPB puzzle games use a different method to localise objects, which requires that the player checks predefined positions in the game field to trigger object sounds. This way, sounds will always appear in a sequence, which relieves the player of having to take in too much simultaneous information. However this also means that the player instead must memorise the positions of objects when playing the game, or frequently check the available positions.

Mudsplat uses a solution that in a sense is a hybrid of the others. In Mudsplat, all objects are associated with brief sounds that are triggered recurrently in a loop. With several sounds triggering repeatedly, each one at a given moment in the loop, no sound risks being "masked" by others. The listener can then obtain a mental overview of the virtual environment by memorising the positions of objects during the intervals when they are silent.

### **Three listening modes**

All interface design deals with the establishment of agreements between the designer and the user. This process gets easier over time, as conventions develop and several different applications begin to use similar interfaces. Today, most sighted people can interpret numerous pictograms, symbols and icons in computer applications. Usually, this communication occurs so effortlessly that we do not think about the underlying agreements that these simplified images are based on. However, when it comes to auditory interfaces, there is a very limited iconography that sound designers can build upon. We can generally recognize simple signals, such as car horns, bicycle bells and telephones, however there is no auditory counterpart to more complex visual languages, such as that of road signs. This is probably why some designers prefer a flexible "earcon" solution, designing an entirely new set of sounds for every new application.

No matter what design approach one chooses, there is always the issue that people tend to listen to sounds in very different ways. For example, some people never listen to the lyrics of songs, while others focus so much on the text that they hardly notice the melody. Another example of this is that many people never notice the onomatopoeic, or iconic, character of words such as "hiccup," "mumble" and "zipper" until it is pointed out to them. A tool for understanding differences in listening can be found in the theory of "listening modes" by Schaeffer [14] and Chion [15]. Chion divides human listening into three modes: causal listening, semantic listening and reduced listening.

Causal listening is used when listening to the source of a sound, attempting to understand what caused it. Gaver refers to this mode of listening as "everyday listening" [16], and points out that very little is known of the cognitive processes

that make us interpret some sounds as “liquid,” “metallic” or “wooden,” and so on. Whether we listen consciously or subconsciously, we can be more or less successful in causal listening depending on our previous experiences with the sound we analyse. This is well illustrated in a study by Alexandersson [17], who found that control operators at a chemical factory, when listening to sound recordings made at the factory, can identify individual machines and determine if a transport band runs too slowly or too quickly or if a bearing is worn out.

We use semantic listening when trying to interpret auditory codes such as speech or Morse code. Just like with causal listening, extensive previous knowledge of the code is required to successfully understand the semantic contents of a sound sequence. Semantic listening to language is traditionally explored within linguistics. In an auditory interface context, semantic listening is used whenever the user wonders “what does this sound mean?” assuming that the application developers have included the sound for a certain purpose.

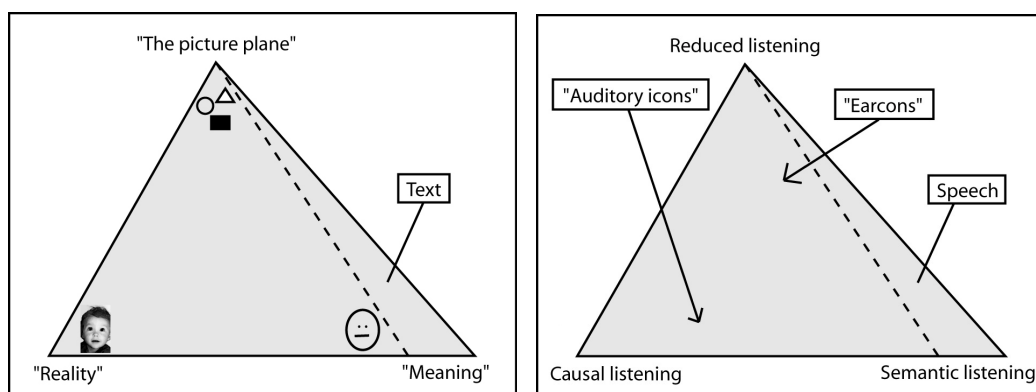
Reduced listening is used when listening to qualities of a sound without considering its source. This is often done when listening for pitches, harmonies and rhythms in music. However, as pioneered by Schaeffer [14], reduced listening can be applied to all sounds with some practice, to appreciate their aesthetic qualities. Gaver uses the term “musical listening” to describe this phenomenon [16]. In an auditory interface context, I think that the term “reduced listening” is more precise, since listening to music often can include causal listening, e.g. when listening to instrumentation, or semantic listening when listening for a narrative in program music. Reduced listening can to a limited extent be described in psychoacoustic or musical terms. However, there is still a very limited terminology for expressing important qualities of sound such as timbre [14] [16] [18].

### **The listener’s perspective**

Even though it is possible to use any listening mode when listening to a sound, it seems like some kinds of sounds encourage certain ways of listening more than others. For example, it takes a lot of effort to apply reduced listening when hearing a spoken sentence in a familiar language. Likewise, it is not common to listen for hidden codes in non-speech sounds, such as birdsong. A better

understanding of how different auditory qualities affect people's ways of listening would be very useful for auditory interface designers.

Sound designers can try to target a certain way of listening, although there is no way to make sure that the listener will comply. In terms of listening modes, the "earcon" approach thus targets semantic-reduced listening, while "auditory icons" target semantic-causal listening. Gaver points out that there are many possibilities to combine the "auditory icon" and "earcon" design approaches [8]. Based on Chion's listening modes, Friberg and Gärdenfors [7] use a triangular model to illustrate the continuum between the three listening modes, emphasising that sounds can be listened to in composite ways. A similar model was originally suggested by McCloud to categorise visual cartoon styles [9] (see Figure 2).



**Figure 2.** McCloud's triangular model of cartoon styles and a corresponding model based on three listening modes. Graphics courtesy of Friberg and Gärdenfors [7], after McCloud [9].

However, McCloud never mentions that the viewer could interpret an image in a different way than the visual artist intended. I believe that this kind of model should refer to ways to interpret images, not the images per se. Designers can target certain modes of looking or listening, although there are always many possible ways to experience the results. In visual art, this fact is well explored by "impossible reality" artists such as Escher [19].

As a design tool, I believe that this model of listening modes can help designers to structure their ideas by targeting certain "fields" of the triangle, sometimes aiming for natural-sounding environments, at other times for more cartoonised or abstract soundscapes. For example, casual listening can be targeted when

designing individual sound objects, while on a larger scale, sounds can be combined and sequenced in a way that emphasises reduced listening.

While sound designers can never be sure that the listener will experience sounds as intended, there are still ways to increase the chances of successful communication. One way to do so is by adding redundancy, which in a sense can be seen as targeting more than one listening mode simultaneously. For example, it can be possible to use an iconic timbre in an alert signal based on a musical phrase that enhances the message. Sounds with many information layers are often used in film and radio plays. For example, a scream can indicate fear or pain, and at the same time reveal the gender and age group of the person screaming, as well as his or her approximate location. Furthermore, the scream can be mixed with, or accompanied by other sounds, such as comical or scary music, to enhance its emotional effect.

Finally, ambiguity should not necessarily be avoided, as it can be an appealing feature in radio, films or in computer games. A classic example of a game that plays with ambiguous listening modes is *Myst* [20], where any visual or auditory detail, which at first seems purely ornamental, can be an important clue to solving puzzles.

### **Conclusions**

Several interesting issues arise when developing entirely sound-based interactive media. Sounds need to be positioned within a virtual soundscape in a way that allows the user to create an accurate mental image of the present objects. Designers must decide what sounds are best suited to represent objects and events within the game, trying to predict how listeners will perceive them. New research in this field could be very useful for designers of auditory interfaces. The lack of an established auditory iconography is an obvious limitation both for designers and users, since sounds often have to be designed and thus learnt from scratch for every new application. Today operating systems for mobile phones and PCs, e.g. Windows XP, offer a variety of sounds that can be assigned to numerous events. However, since the user is allowed to completely customise the auditory interface using any sound to indicate any type of event, this does



little towards the establishment of a standard. Hopefully this problem will fade away as auditory interfaces mature and become more widely used.

## References

- [1] Audiogames.net. [homepage on the Internet] <http://www.audiogames.net/> (Accessed 2006-08-15).
- [2] W. W. Gaver. Auditory interfaces. In M. G. Helander, T. K. Landauer, and P. V. Prabhu, editors, *Handbook of Human-Computer Interaction*, 2nd ed., Elsevier, Amsterdam, 1997, pp. 1003–1041.
- [3] Talboks- och Punktskriftsbiblioteket. [homepage on the Internet] [http://www.tpb.se/barnens\\_tpb/spel/ljud/](http://www.tpb.se/barnens_tpb/spel/ljud/) (Accessed 2006-08-15).
- [4] Timgames.org. [homepage on the Internet] <http://www.timgames.org/mudsplat> (Accessed 2006-08-15).
- [5] Y. Eriksson and D. Gärdenfors. Computer games for children with visual impairments. In *Proceedings of the 5th International Conference of Disability, Virtual Reality & Associated Technologies*. Oxford, UK, 2004, pp. 79-86.
- [6] D. Gärdenfors. Designing sound-based computer games. *Digital Creativity*, 14(2), 2003, pp. 111–114.
- [7] J. Friberg and D. Gärdenfors. Audio games: new perspectives on game audio. In *Proceedings of the Advances in Computer Entertainment Technology'04*, Singapore 2004, pp. 148-154.
- [8] W. W. Gaver. The Sonic Finder: an interface that uses auditory icons. *Human Computer Interaction*, 4, 1989, pp. 67-94.
- [9] S. McCloud. *Understanding comics: the invisible art*. Harper Collins Publishers, Inc, New York, 1993.
- [10] C. Stalling. *The Carl Stalling project: music from Warner Bros. Cartoons 1936-1958*. [Music CD and leaflet] Warner Bros. Records, 1990.
- [11] M. Blattner, D Sumikawa and R. Greenberg. Earcons and icons: Their structure and common design principles. *Human-Computer Interaction* 4(1), 1989, pp. 11-44.
- [12] A. Menshikov. *Modern audio technologies in games*. [article on the Internet] <http://www.digit-life.com/articles2/sound-technology/> (Accessed 2006-08-15).
- [13] F. Winberg and S.-O. Hellström. The quest for auditory direct manipulation: the sonified Towers of Hanoi. In *Proceedings of the 3rd International Conference on Disability, Virtual Reality and Associated Technologies*, Alghero, Italy, 2000, pp. 75–81.
- [14] P. Schaeffer. *Traité des objets musicaux*. Editions du Seuil, Paris, 1966.

[15] M. Chion. *Audio-vision: sound on screen*. Gorbman C, editor. Columbia University Press, New York, 1994.

[16] W. W. Gaver. What in the world do we hear? An ecological approach to auditory event perception. *Ecological Psychology*, 5(1) 1993, pp. 1–29.

[17] P. Alexandersson. Peripheral awareness and smooth notification: the use of natural sounds in process control work. In *NordCHI 2004*, Tampere Finland 2004, pp. 281-284.

[18] T. Wishart. *On Sonic Art*. Revised ed. Contemporary Music Studies Series vol. 12. S. Emmerson editor. Harwood Academic, Amsterdam, 1996.

[19] M. C. Escher. *The graphic work of M. C. Escher: introduced and explained by the artist*. Pan Books, London, 1975.

[20] Cyan Worlds [homepage on the Internet] <http://www.cyanworlds.com/> (Accessed 2006-08-15).

### **Abstract**

When developing computer games for visually impaired users, it is possible to create complex interactive environments using sounds only. This paper points out key issues involved when designing games based on "auditory interfaces". A multitude of different types of sounds can be used, including recorded or synthesised speech, realistic sound effects or abstract, musical sounds. Sounds must be positioned in a "virtual soundscape" and the listener must be able to obtain a mental overview of the game space. Three ways of listening are introduced, emphasising the possibilities to design sounds that carry information on many levels simultaneously.

### **Om författaren**

Dan Gärdenfors har en magisterexamen i musikvetenskap från Lunds Universitet och har sedan 1999 intresserat sig för ljud i interaktiva medier. Han har utvecklat datorspel för svenska och danska synskadade, arrangerat utställningar med ljudkonst och planerat en utbildning inom interaktivt ljud i Hässleholm. I tre år jobbade han med forskning och utveckling av ljudspel på SITREC (Stockholm International Toy Research Centre) inom ett EU-finansierat projekt. Sedan 2005 arbetar han som interaktionsdesigner på Anoto, ett Lundabaserat företag som utvecklar digitala pennor.

## **OPERATIONAL SOUNDS – THE FACTORY SOUNDSCAPE AS CARRIER OF INFORMATION**

Petter Alexanderson, Konrad Tollmar  
Department of Informatics, HCID Group  
Ole Römers Väg 6, S-22363 Lund  
petter.alexanderson@ics.lu.se

In this text we present a vision for a flexible and configurable auditory environment. In particular, we would like to bring forward design principles to support everyday soundscapes that enable informal interaction, e.g. for notification and awareness of the surrounding world.

The starting point is a phenomenological study of the auditory environment in a chemical factory, and how a group of process operators ascribe meaning to sound in their daily work environment. Designing auditory interfaces that will harmonize with existing work practices require an understanding of the acoustic ecology, i.e. the relation between people and the existing soundscape 'in use'. Based on our findings we have explored several design directions for a sound-augmented work practice. Three design directions will be discussed in this text which stem from the daily work of the process operators in the 'C-shift' at a chemical factory.

### **Exploring sound and soundscapes**

A series of visits were made to the factory in order to capture the factory soundscape. The operators were followed during daily work, and the sounds of the factory were recorded along with corresponding video material. A selection of 17 sound clips was used to produce audio CD's from the factory. During Participatory Soundscape Interpretation (PSI) sessions, six operators using individual head phones listened to the CD's. Their thoughts about each clip were discussed. The operators were asked to 1) Describe what they heard, 2) Indicate from which location each clip originated, 3) Describe what was going on, and 4) Any other associations made while listening. After

each sound clip, the corresponding video material was watched, and the PSI sessions were in turn recorded on video for further analysis. The analysis inductively derived a structure of themes and categories based on the concepts and classifications used by the operators themselves. A phenomenological approach was preferred in order to illustrate the meaning of sounds as described by the operators themselves, in order to capture information not fitting into a predefined classification scheme.

### **Understanding the auditory environment**

From PSI session transcripts, several themes relating to the daily use and interpretation of environmental factory sounds emerged. Here, three aspects of the soundscape meaning are presented – namely how the operators use sound for identifying things and places, notification about status of surrounding artifacts, and maintaining social awareness.

#### *Objects and Places*

In the factory control room, there are several printers, both line printers and laser printers and most of them have a dedicated function. Operators could clearly identify which printer was printing and what kind of output they produced. The line printers sound about the same, but different printers print a different number of rows, so it is apparent which device is active and what is about to be printed.

#### *Notification of Status*

One of the few remaining manual operations in the control room is the batching of raw material into different reactors. The rest of the process control is to a large extent automatic and is monitored and controlled by computer systems. The batching commences when the level of liquid in certain tanks reaches a specified level. The operators must constantly monitor the process and start the batch at the exact right moment. Lamps on a panel and tank levels in the computerized process control system indicate when it is time. However, the operators must give attention to visually monitoring the lamps and tank levels. Another, more subtle but precise level indicator is the sound of the relay clicks from the wall panel informing the operator when the batching should start.

### *Social Awareness*

The PSI sessions at the factory show how sounds can provide valuable information to operators about their colleagues' activities and location. This includes information about what someone is doing, where in the plant the activity is taking place and, in some instances, also who is performing a particular function. In the control room the sound of someone turning the wheel of a line printer to feed paper, someone using a keyboard to sign an alarm in the computer system or someone pressing the wrong key tells a quite detailed story about what is going on, at least to those closely acquainted with the factory soundscape. Listening to the sound of control room printers, it was possible to determine that someone from another department had printed a document since one of the printers was dedicated to them. The whirling of the coffee machine was another sound event notifying the operator that someone was in the kitchen – in most cases to talk to colleagues or to take a break. There were occasions when the operator listening to a sound clip also could give information about who they were hearing.

*[Laughs] Someone is using the coffee machine...it is filling up coffee in a plastic mug in the control room. Probably it is a guest, perhaps you [referring to one of the authors]. At least it is not one of us since I hear a plastic mug and not a porcelain one.*

(Johan, bag operator)

We have reason to believe that the auditory border resources contribute considerably to guide interaction in daily factory work. The next issue became to explore how this could be utilized in soundscape design.

### **Designing Interactive Soundscapes**

Based on the findings from the case study, we have explored several design directions for auditory environments. An important part of the design process was the process operators' contribution to the suggested concepts. A series of four design workshops were outlined and carried out in cooperation with a group of operators at the factory. The concepts described below should hence be regarded as an attempt at working examples, as well as a means for further exploration.

### *SonicProbe - Collecting soundmarks*

The leading idea of the probing device sketched out here, called a 'SonicProbe', is to elaborate on the natural character of sound as a medium, that includes limitations related to time and space. Sounds have a temporal quality, which means that once they cease, they must have been recorded in order to be experienced again. Overcoming space limitations involves enhancing sounds so that they propagate outside their space and place of origin. In this way the SonicProbe adds auditory persistence and scope – to objects as well as to rooms.

The SonicProbe was designed to be a personal device for gathering everyday sounds, and thereby to address inherent limitations of the auditory medium (image 2). If persistence and mobility are one side of the coin, then the other is the idea of making sounds common goods by making them public.

### *SonicRep - Reviewing and sharing*

The 'SonicRep' concept elaborates on how to interact with sounds originating from a single source but from different points in time. In a way, the SonicRep could be compared to a photo album, in which sounds collected from different times and places could be gathered and made accessible (image 3). Through the SonicProbe and SonicRep concepts, we aim to enhance objects and rooms by offering new ways to collect and access sounds over time, and to augment people by extending the human ability to review sounds by making sounds from different points of time easily accessible. This will also facilitate the sharing of auditory experiences.

### *ScapeNav – Sonic 'hot spots'*

The third direction addresses the issue of how to remain aware of multiple ongoing things of interest, and how to 'stay tuned' without being drowned in concurrent impressions. This is undoubtedly a contemporary issue as information technology is pervasive in our society and can create rather demanding and stressful situations when it comes to attention and cognitive load. The contribution here is a real-time 'awareness mechanism' and an exploration into the design of a sound-based awareness environment. Although the examples in the text are from the factory, we present the 'ScapeNav' concept as a general idea supporting the desire to 'keep up to date' in the realm of everyday life (image 4). In the case presented here, machines will

be used as examples, because the concept was developed with process operators as committed participants. The augmentation approach here is 1) to augment the personal space with facilities for enhanced remote awareness of 'hot spots' and with an increased capacity for parallel listening, and 2) to mediate the sound of remote places and objects, bringing them together in a superimposed soundscape.

The ScapeNav is also a personal device, which mediates the dynamics of spatial focus points, what we refer to as 'hot spots', in real-time. The following matters were in focus for the design concept:

- the dynamic interactive configuration of 'hot spots'
- the issues of perceiving complex soundscapes
- elaboration with interface concepts for a superimposed soundscape

The three design directions were revisited with reference to the daily work of the process operators in the C-shift at the factory. Contextualizing the concepts for soundscape augmentation is a vital issue in order to give the ideas a 'stage' on which they can be acted out, and a means for a practical 'try out'. This exploration has provided us with new and valuable insights, and hence is a vital part of the design process.

## **Evaluating Interactive Soundscapes**

### *The SonicProbe – Mobile recording*

As an exploration of what it means to collect valuable sounds, an informal try-out was carried out by Bertil, one of the control room operators. Using a product from the consumer market, Bertil carried the device in his chest pocket for a period of three weeks. The aim of the test was twofold: to see what qualities a portable sound recording device would need to have, and also to use the device as a tool for finding, recording and selecting a number of sonic hot spots. These findings would then be used as references and working material for the design project. The collecting of factory sounds hence became a hands-on way of 'probing' the idea of a sonic probe. The sound recorder in itself had appealing qualities in several ways: it was portable,

small and light, and a device with few controls, which made it fairly straightforward to use.

#### *The SonicRep - A shared repository*

During the workshops, a recurring discussion centered around how to use sound references for problem solving and problem prevention. There was some ambivalence about one issue, namely how to arrange for collective listening. For example, in order to provide for discussions and negotiations about the sound of a bag machine, there is a need for group listening in an undisturbed environment. However, colleagues working in adjacent areas of the control room could find this annoying or distracting. 'Letting loose' the factory sounds in the control room, could result in an uncontrollable and unbearable work environment. The use of headphones was considered, but was abandoned because of the obvious risk of losing the real-time reference which simultaneous listening provides.

#### *The ScapeNav – A superimposed soundscape*

The ScapeNav concept was discussed with the C-shift operators during an early stage of the design project. The idea of creating a superimposed soundscape of the factory was tried out and reviewed. Now, the dynamic configuration of hot spots and options about how to perceive a complex soundscape were on the agenda. An interactive prototype was prepared with sounds collected and selected from the hot spots. A major limitation of the prototype, with regard to using a ScapeNav in practice, was that the sounds from the hot spots were pre-recorded – not picked up in real-time. A consequence of this was that certain dynamics, described by the operators, were never captured on the recordings. However, in cases where the lack of dynamics became critical, a change in a sound was simulated using audio editing software. For the trials, the prototype was installed both on a laptop computer and in a handheld device. The prototype was first tested and reviewed by the C-shift informally. After a short presentation, and a period of time to become familiar with the prototype, the operators were asked to envision how the device would be used by them during the work day. They were also encouraged to 'show and tell' how they would use the ScapeNav out on the factory floor.



## **Taking it further**

The collection and joint listening of everyday sounds also has meaning outside the factory context – consider a recording of the sound of a malfunctioning car engine, or the first cry of a newborn child. We also argue that the history of sound events has importance in everyday life, and so the SonicProbe and SonicRep are useful and valuable concepts not only for factory workers.

The process control examples show us the value of awareness services. However, one could still argue that this is too narrow a case for extrapolation. On the contrary, we maintain that this is a way to understand the everyday organization, not only of process control work, but also of daily life. Domestic environments are undoubtedly socially organized, and collaboration, cooperation, and coordination structures are obviously present. In the everyday organization of life, a number of ‘awareness structures’ are interwoven in social practice. The rationale for awareness displays is that everyday interaction can be characterized to a great extent as a matter of coordination and ‘keeping up to date’, e.g. keeping an eye on the Sunday roast, listening/becoming aware of the kids safe return home, or making sure that the sailboat remains above water at the summer cottage.

Finally, we argue for a phenomenological approach to understanding auditory interfaces and soundscapes for cooperative environments. Existing work practices offer a wealth of information. Valuable information can be gleaned from how already present sounds are used, and by letting this knowledge inform the design of new useful auditory environments.

## **Acknowledgments**

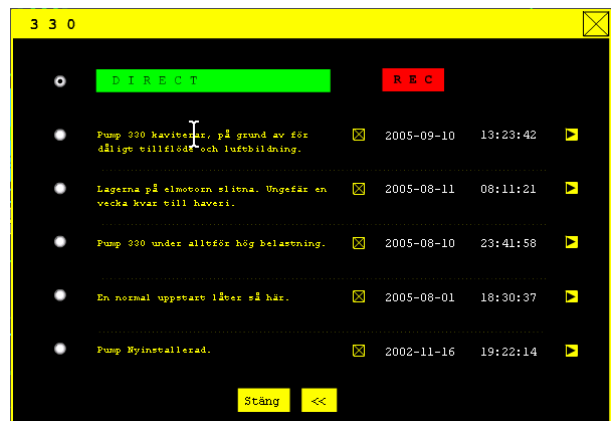
We thank all reviewers who wrote and provided helpful comments on previous versions of this text. Last but not least a special thanks to Bertil Cronholm and Börje Klemmedsson in the factory C-shift.



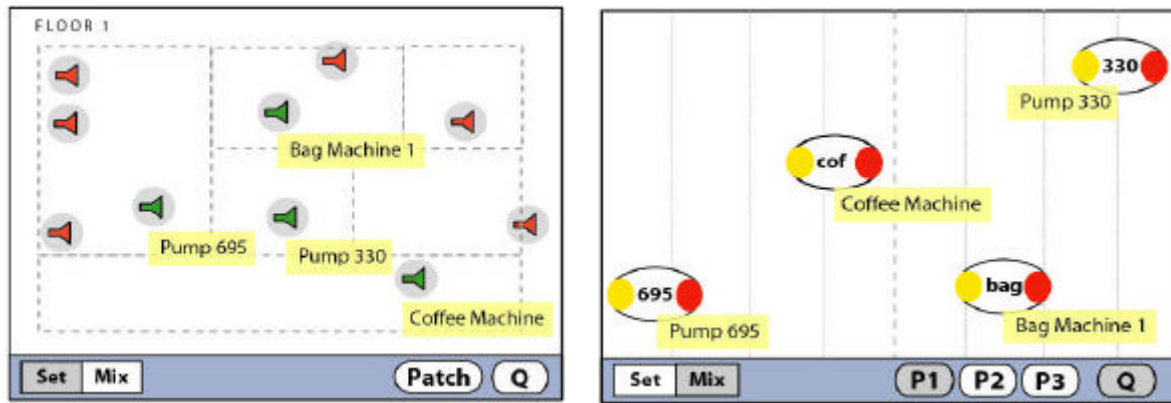
**Image 1:** Factory snapshots from the bag hall, the centrifuge area, and the quality test area.



**Image 2:** A SonicProbe foam mock-up and the probing device. The probes has a built-in 1.3 MegaPixel camera, and changeable microphones to allow for uni- as well as omnidirectional recordings.



**Image 3:** A SonicRep prototype. Here the audio history for the '330' pump is made available as a shared resource for the operators. The list of recordings could be modulated, and new sounds added via a recording function.



**Image 4:** The ScapeNav prototype. Left: the 'set'-screen used to toggle sound sources on/off. Right: the 'mix' screen used to set pan and volume for each sound source. The 'patch' option enables the change of sound representation for a specific sound source.



**Image 5:** The ScapeNav prototype 'out-in-the-wild'. Bertil also wears a bluetooth-equipped headset for wireless audio.

## Icke-visuella audio-haptiska datorgränssnitt

Kirsten Rasmus-Gröhn, Certec, Inst för Designvetenskaper,  
Lunds Universitet



Certecs forskning om innovativa datorgränssnitt för personer som är blinda eller har andra synskador är baserat på en kombination av ljudåterkoppling och känselåterkoppling (haptik). Vårt mål är att kunna göra datorburen information tillgänglig som hittills inte är tillgänglig för blinda. Det kan t ex gälla bilder, kartor eller virtuella världar. Om man inte kan se måste man i så fall utnyttja t ex känsel och hörsel på ett kreativt sätt för att dela med sig av sådan information. Vi använder oss av en kombination av olika sorters ljud och speciell känsel-återkoppling för att illustrera t ex en karta.

För känsel-återkopplingen används "The PHANToM", en liten robotarm som man håller i och som kan motverka handens rörelse. På så sätt kan man uppleva att man känner på föremål som i verkligheten endast finns som en datormodell. För ljudåterkoppling använder vi oftast ljud som spelas upp i hörlurar, men det kan också vara vanliga datorhögtalare.

För närvarande arbetar vi främst med två olika tillämpningar: ett audio-haptiskt ritprogram och en audio-haptisk trafikmiljö i 3D.

### Audio-haptiskt ritprogram

Ritprogrammet gör det möjligt för en blind användare att rita med en virtuell penna på ett virtuellt papper. De linjer som användaren ritar går också att känna som en relief. Det finns även verktyg för att hjälpa användaren att rita t ex exakta cirklar och rektanglar, samt att hantera ritade objekt (förstora, förminska, flytta, kopiera, klistra in och ta bort). För att användningen ska vara enkel och det ska vara möjligt att förstå vad som händer



är ljuddesignen planerad så att man ska kunna få ljudinformation på olika sätt:

- ✍ Pennans position kan höras genom tonändring:
  - Frekvens (upp & ner) 200Hz – 1kHz
  - Linjär stereo (vänster – höger)
- ✍ Talad information om objekt (vilken typ av objekt som är aktivt)
- ✍ Händelsestyrda ljudeffekter vid objekthantering
  - Storleksförändring, flyttning, form-ändring
  - Ta bort, kopiera, klistra in
- ✍ Ljudsignal för att känna igen kanterna på arbetsområdet

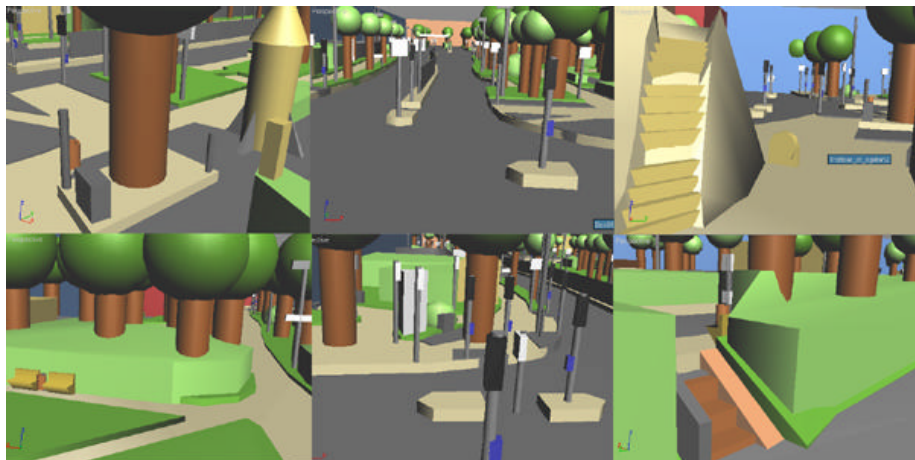
### Kommande arbete på ljudåterkoppling i ritprogrammet:

- ✎ Experimentera med ljudåterkoppling för att förbättra positionsinformation
- ✎ Göra det möjligt för användaren att sätta ut ljudfyrrar
- ✎ Lägga till olika nivåer av ljudinformation (nybörjare → expert)

### Audio-haptisk virtuell trafikmiljö

Den virtuella trafikmiljön gör det möjligt för en blind användare att prova och öva i en ofarlig modell av en trafikmiljö innan han eller hon går ut i den riktiga miljön. För att kunna uppleva världen och känna på den används även i detta fall "The PHANToM" för känselåterkopplingen. För ändamålet har vi tagit fram en 3D värld, en modell av trafikområdet kring Mejeriet i Lund, med 484 objekt, både statiska och dynamiska. De flesta objekt är statiska, såsom träd, lyktstolpar, bänkar, refuger, buskar osv. Bilar och cyklar är dynamiska genom att de rör på sig. Trafikljusen är dynamiska på så sätt att de styr trafiken – och därmed också ger ljudsignaler vid övergångsställena. Vidare finns ljudinformation om:

- ✎ Talad information om objekt
- ✎ 3D-ljud på bilar och övergångsställen med ljudsignal
- ✎ Händelsestyrda ljudeffekter
  - Om avataren blir påkörd spelas ljudeffekt
  - Ljudsignal för att känna igen kanterna på arbetsområdet
- ✎ Ljudeffekter för interaktion med miljön
  - Scrollning av arbetsområde
  - Skalning av arbetsområde



### Läsa mer

*Enabling Audio-Haptics* Kirsten Rasmus-Gröhn, Licentiatavhandling, september 2006, [www.certec.lth.se/doc/enablingaudiohaptics](http://www.certec.lth.se/doc/enablingaudiohaptics)

*A Virtual Traffic Environment for People with Visual Impairments*

Charlotte Magnusson, Kirsten Rasmus-Gröhn

Visual Impairment Research, pp 1-12 Vol 7. No 1. 2005

*Slutrapport – Projekt haptik och trafik - en förstudie, beteckning EK 50A 2002:5083*

Charlotte Magnusson Internrapport och rapport till uppdragsgivaren Skyltfonden

[www.certec.lth.se/haptics](http://www.certec.lth.se/haptics)











# LUNDS UNIVERSITET

---

Lyssnande Lund - Ljudmiljöcentrum vid Lunds universitet

Box 117, 221 00 Lund  
Telefon 046-222 09 46.

[www.ljudcentrum.lu.se](http://www.ljudcentrum.lu.se)