



**Linköpings universitet**  
TEKNISKA HÖGSKOLAN

# **Röststyrning av aggregate på en engreppsskördare En Wizard of Oz-studie**

Martin Englund

Examensarbete  
Institutionen för ekonomisk och industriell utveckling  
LIU-IEI-TEK-A--09/00666—SE

## Sammanfattning

Syftet med denna studie var att utreda hur röststyrning kan användas för kontroll av aggregatet på en engreppsskördare.

För att studera röststyrning genomfördes en så kallad Wizard of Oz-studie. Metoden går ut på att en människa ersätter en funktion i ett system som ännu inte finns eller är färdigutvecklad. I studien var det röstigenkänningen och inmatningen av korrekt respons som utfördes av en människa, trollkarlen. Rollen som trollkarl utfördes av en före detta skördarförare som numera arbetar som forskare på Skogforsk.

Studien genomfördes i Skogforsks skogsmaskinsimulator. I den kan en försöksperson på ett verklighetstroget sätt arbeta med en engreppsskördare. Som försökspersoner användes fem professionella förare.

Innan en förare provade röststyrningen fick han köra simulatören med konventionell styrning för att vänja sig vid hur maskinen kändes att köra i den simulerade miljön. Därefter fick föraren använda sig av röststyrning. Försökspersonerna som var omedvetna om trollkarlens existens fick endast instruktion om att slutavverka ett bestånd genom att kontrollera aggregatet med rösten. De fick inga förslag på kommandon eller uttryck att använda utan fick själva utforska systemet och använda det som de ville. Kranen kontrollerade de som vanligt med spakarna.

Under studien observerades försökspersonernas agerande av försöksledaren och efteråt genomfördes en intervju om upplevelsen av röststyrningen.

Alla förarna lärde sig snabbt att använda systemet och använde sig av samma princip. En knapptryckning ersattes med ett kommando. Förarna använde sig till en början av uttryck de var bekanta med sedan tidigare. De rationaliserade medvetet uttrycken de använde till att bli korta och nästan alltid bara innehålla ett ord.

Överlag tyckte försökspersonerna att röststyrningen var en positiv upplevelse och tyckte att det fungerade bra men förutsåg att det skulle bli problematiskt att använda vid svårarbetad skog som kräver mycket manuell inblandning i apteringen.

Röststyrningen verkar gynna användning av automatiska sekvenser.

## **Förord**

Den här rapporten är resultatet av mitt examensarbete på civilingenjörsprogrammet i maskinteknik med inriktning mot ergonomidesign vid Linköpings tekniska högskola.

Arbetet genomfördes hos Skogforsk i Uppsala i samarbete med avdelningen för industriell arbetsvetenskap, IAV, vid Institutionen för ekonomisk och industriell utveckling, IEI, vid LiTH.

Jag vill tacka Björn Löfgren som var min handledare på Skogforsk och Torbjörn Alm och Kjell Ohlsson som var mina handledare vid LiTH.

Tack till Petrus Jönsson som var min trollkarl och alla andra på Skogforsk som varit mig behjälpliga.

Tack till Andreas Lind och alla andra på Oryx Simulations som hjälpt mig med simulatorm.

Sist men inte minst ett stort tack till alla er som ställde upp som deltagare i mitt försök.

/Martin Englund

## Innehåll

Sammanfattning .....	1
Förord.....	2
Innehåll .....	3
Inledning .....	5
Skogforsk .....	5
Bakgrund.....	5
Tvåmaskinsystem .....	5
Skogforsks simulator .....	6
Syfte.....	7
Begränsningar .....	7
Teori.....	7
Arbetsmiljön i skogsmaskiner .....	7
Röststyrning .....	8
Igenkänningsfel .....	9
Svårigheter med talljud .....	9
Indirekta miljöbetingade faktorer .....	9
Direkta miljöbetingade faktorer .....	10
Uppmärksamhet .....	10
Felhantering .....	10
Feedback .....	11
Metod.....	12
Studie i simulator .....	12
Observation .....	12
Riktad öppen intervju .....	12
Wizard of Oz.....	13
The Wizard - Trollkarlen.....	13
Dialogen .....	14
Uppgiften .....	14
Försökspersonerna.....	15
Intervju och debriefing .....	15
Genomförande.....	16
Fältobservation och intervju .....	16
Wizard of Oz-simulering.....	16
Trollkarlens arbetsplats .....	16
Skogen.....	20
Försökspersoner .....	21
Inställningar av maskinen i simulatorn .....	21
Introduktion.....	22
Övning.....	22
Presentation av röststyrningsstudien .....	22
Uppgiften .....	22
Körning med röststyrning.....	22
Intervju.....	23
Resultat .....	23
Övning.....	23
Körning med röststyrning.....	23
Observation och intervju .....	23
Diskussion.....	26

Simulatorn.....	26
Intervjun.....	26
Wizard of Oz-metoden.....	27
Inspelad data .....	27
Framtida forskning.....	27
Slutsatser .....	28
Referenser .....	30

# Inledning

## Skogforsk

Skogforsk är det svenska skogsbrukets forskningsinstitut. Det sysselsätter för närvarande runt ett hundra personer varav ett sextiotal är forskare. De sysslar med tillämpad forskning inom bland annat skogsträdsförädling, skogsskötsel och miljö, driftsteknik råvaruutnyttjande och logistik. Meningen är att forskningen ska komma skogsbruket till nytta och stora resurser läggs på att förmedla och underlätta implementeringen av forskningsresultaten.

Basen för verksamheten ligger i Uppsala. Där finns de flesta forskarna samt forskningsledning, administration och informationsavdelning. Vid Skogforsks två forskningsstationer, Ekebo i söder och Sävar i norr, arbetar man bland annat med förädling och plantforskning.

Skogforsk finansieras till hälften av skogsnäringen och till andra hälften av staten. Totalt omsätts omkring 100 miljoner kronor. (<http://www.skogforsk.se>)

## Bakgrund

### Tvåmaskinsystem

Skogsavverkningen är idag ett mekaniserat arbete. Man använder sig av ett tvåmaskinsystem bestående av en engreppsskördare och en skotare.

Engreppsskördaren är en terränggående maskin som fäller och upparbetar trädet. Den har en ledad hydraulisk kran som längst ut bär på aggregatet som upparbetar eller apterar trädet. Föraren för med den hydrauliska kranen fram aggregatet till det träd han tänker fälla och griper tag i det med aggregatets kvistknivar och matarhjul. Svärdet som sitter längst ner på aggregatet kapar trädet och föraren fäller trädet i lämplig riktning. Trädet förs sen genom aggregatet av matarhjulen och kvistknivarna slår då av grenarna. När en lämplig längd matats fram kapas den till en stock som läggs på hög. Vad som är en lämplig längd avgörs i de flesta fall av en dator, som med information som aggregatet ger om trädets tjocklek och frammatad längd, samt aktuella prislistor väljer den längd som är mest lönsam. Föraren ingriper och ändrar datorns beslut om det förekommer några avvikelser som till exempel att stammen är krökt eller drabbad av röta. Beroende på kvalitet och diameter lägger föraren upp stockarna i högar av grovtimmer, klentimmer, massaved eller något annat sortiment.

Skotaren följer i skördarens spår för att samla upp stockarna. Den har liksom skördaren en hydraulisk kran men på den sitter endast en gripklo. Med dess hjälp samlar skotarföraren upp stockarna och lägger dem i lastutrymmet. Skotaren för sedan ut stockarna till en uppsamlingsplats där de kan nås av en timmerbil.

Ett alternativ till tvåmaskinsystemet är drivaren. Den är en maskin som har både skördarens och skotarens funktion, dvs. den har både ett aggregat för fällning och upparbetning av stockarna, men också ett lastutrymme för transport. Detta system används dock inte alls i lika stor utsträckning som tvåmaskinsystemet.

Styrningen av maskinerna sker idag med ett tvåspakssystem. På eller i närheten av spakarna finns en mängd knappar för att kontrollera alla funktioner. Kontrollerna innebär fler knappar hos skördaren, eftersom den har en uppgift som innebär många fler moment än skotningen.

Arbetet har ett högt tempo och innebär en stor mental belastning på maskinförarna. Speciellt gäller detta för skördarförarna. Förarna måste hela tiden fatta kvalificerade beslut. Tvåspakssystemet innebär en statisk arbetsställning som kan leda till arbetsskador. Det finns därmed anledning att försöka förbättra skogsmaskinförarnas arbetsmiljö för att minska den fysiska och mentala arbetsbelastningen.

### **Skogforsks simulator**

Skogforsks simulator är en skogsmaskinsimulator utvecklad av Oryx Simulations. I den är det möjligt att simulera användning av alla typer av skogsmaskiner. Simulatorn har tre skärmar som omger föraren. I mitten av de tre skärmarna står en riktig förarstol med kontrollspakarna monterade på armstöden. Skogen där scenariot utspelar sig projiceras på skärmarna av tre projektorer, som sitter utanför ”hytten”. Eftersom denna studie handlade om skördare användes en stol från en skördare.



*Simulatorn.*

## Syfte

Undersökningens syfte var att utvärdera hur förare av en engreppsskördare kan använda sig av naturligt tal för styrning av aggregatet.

## Begränsningar

Studien begränsades av tillgänglig tid. Metoden skulle kunna användas iterativt genom att för varje iteration lägga till nya lärdomar till utvecklingen av ett röststyrningssystem för engreppsskördare.

## Teori

### *Arbetsmiljön i skogsmaskiner*

Skogsavverkningen genomgick stora förändringar under senare hälften av 1900-talet. Från att ha varit ett arbete som sköttes med motorsåg till en situation där nästan all avverkning sker med ett system av engreppsskördare och skotare. Tidigare var arbetet tungt och belastade hela kroppen, men idag innebär det ett stillasittande förararbete med små högrepetitiva handrörelser. Nästan allt arbete i maskinerna styrs med multifunktionella spakar på vilka det sitter en mängd knappar. [Attebrant m.fl. 1998]



*Spak av den modell som används i studien.*

Framsteg inom driftsäkerhet har gjort att förarna behöver ägna betydligt mindre tid åt underhåll av maskinerna och mer sällan har någon anledning att vara ute ur förarhytten. I



mitten av 70-talet spenderade en maskinförare 60 procent av arbetsdagen i hytten varav 80 procent var reglagemanövrering. I början av 90-talet hade de siffrorna ökat till 80 respektive 90 procent. Variationen i förarnas arbete har minskat samtidigt som varaktigheten har ökat. Den kombinationen anses särskilt skadlig eftersom samma muskelfibrer utnyttjas hela tiden och det inte finns möjlighet till återhämtning. [Attebrant m.fl. 1998]

Både teknisk och organisatorisk utveckling har bidragit till ändringen av arbetsförhållandena i avverkningsarbetet. Idag låter skogsbolagen entreprenörer sköta största delen av avverkningen istället för egna anställda. Entreprenörerna, som själva måste investera i dyra skogsmaskiner, måste ofta arbeta 50 till 70 timmar i veckan för att få tillräcklig lönsamhet. [Attebrant m.fl. 1998]

Reglagearbetet innebär självklart belastningar i händer och underarmar, men också i nacke och skuldror. [Eklund 1998]

Under 70 och 80-talen gjorde tekniska förbättringar att muskelaktiviteten i skuldran minskade. Studier visar att medelbelastningen på övre delen av trapeziusmuskeln i skuldran vid reglagearbete minskade från 16 till 4 procent av maximal aktivitet mellan 1970 och 1995. Under 90-talet gjordes vidare studier som visade att ytterligare förbättringar i utformningen av reglagen inte gav någon ytterligare reduktion av aktiviteten i skuldran. Orsakerna till muskelaktiviteten har troligen andra förklaringar. De höga kraven på precision och koordination vid reglagemanövreringen kan ge upphov till aktivitet i skuldran. Helkroppsvibrationer som överförs från maskinen fungerar som en förstärkare. En annan orsak till muskelaktivitet i skuldran är de kognitiva och perceptuella krav maskinföraren upplever då han i högt tempo måste planera avverkningen och välja vilket träd som ska fällas näst. [Attebrant m.fl. 1998]

Skogsbranschen var en av de första att utsättas för belastningsergonomisk forskning och 1969 kom den första versionen av Ergonomisk Checklista för Skogsmaskiner. Sedan dess har den uppdaterats i flera omgångar. Checklistan inriktade sig på utformningen av maskinerna. Tillsammans med den tekniska utvecklingen har den troligen bidragit till de förbättringar som skett. [Attebrant m.fl. 1998]

## **Röststyrning**

Genom att kommunicera med maskiner genom tal får människan en psykologisk fördel då kommunikationen sker på våra villkor. Händer och ögon kan lämnas fria för andra uppgifter. [Blomberg & Elenius, 2005]

Automatisk igenkänning av tal har blivit allt mer intressant i takt med att datorerna har utvecklats. Många av de problem som tidigare fanns med tekniken har idag i stor utsträckning övervunnits. Till exempel krävdes förut att ord separerades med en kort paus och att talaren tränade med programmet. Ordförrådets storlek var starkt begränsad och igenkänningsgraden var ändå låg. Idag finns talaroberoende system som utan träning kan känna igen kontinuerligt tal, utan paus mellan orden, i väldigt hög grad (upp till 99 % enligt tillverkarna) och som har ett närmast obegränsat ordförråd. System kan till och med anpassa sig efter hur talarens röst ändras vid t.ex. trötthet eller förkylning. [Blomberg & Elenius, 2005]

Tekniken används idag bland annat för direkt diktering i ordbehandlare. Ett av de områden där detta har funnit stor professionell användning är hos läkare och tandläkare, som kan sköta

dokumentationen av en undersökning under tiden den utförs. Alternativet är att dokumentera i efterhand eller att låta en sekreterare ta anteckningar.

Styrning av till exempel klimatsystem och ljudanläggning i bilar via röstkommando är ett annat användningsområde. Många har också kommit i kontakt med taligenkänning i automatiska tjänster via telefon, till exempel för biljettbokning.

Gellatly (1997) ger riktlinjer för när det kan vara användbart med röststyrning. (1) Tal ska bara användas när insignaler behövs sällan. (2) Tal ska användas konsekvent för en del av en uppgift, t.ex. använd tal för bara kommandon istället för till kommandon och datainmatning. (3) Funktioner inom en uppgift som styrs med tal ska också väljas med omsorg. Till exempel är det inte lämpligt att använda tal till funktioner som att flytta en muspekare på en skärm eller justera backspeglarna i en bil.

## Igenkänningsfel

Ett av de största problemområdena med taligenkänningsystem är igenkänningsgraden. Enligt Gellatly (1997) finns det tre skilda kategorier fel som kan ske i igenkänningen.

Den första kategorin är substitutionsfel. Det sker när ett ord som finns med i systemet känns igen som ett annat ord som också finns med. Till exempel om ordet "träd" uttalas, men systemet känner igen uttalandet som ordet "skog" istället.

Den andra kategorin är förkastande fel. Det inträffar då ett ord som finns med i systemet uttalas men inte känns igen. Till exempel uttalas "träd" men systemet reagerar inte på något sätt.

Den tredje kategorin fel är instoppningsfel och inträffar då systemet tolkar ett ljud som ett ord som finns i systemet. Till exempel skulle systemet kunna känna igen ordet "skog" då störande ljud förekommit i närheten av mikrofonen trots att ordet inte använts.

Användare har minst tolerans för substitutionsfel och instoppningsfel. Anledningen är att i de fallen måste användaren rätta till det fel som uppstått. Vid ett förkastande fel behöver ju användaren förhoppningsvis bara upprepa ordet. [Gellatly, 1997]

## Svårigheter med talljud

Talet är en komplicerad ström av ljud. Normalt finns det inget uppehåll mellan ord som det finns i text. Talorganets mekaniska egenskaper gör att talet blir en kontinuerlig process. Orden sitter ihop med varandra och uttalet av ett ord påverkas av vilka ord som kommer före och efter. Ett ord som står i fokus i en mening uttalas ofta tydligt, medan till exempel ändelser, artiklar och prepositioner uttalas otydligt eller helt enkelt utelämnas. Anledningen till att människor ändå kan förstå varandra är att vi har djup insikt och erfarenhet av språkets natur. Vi använder oss av kunskap om talaren, ämnet och sammanhanget för att lägga till information. På så sätt kompenserar vi för information i talsignalen som förvanskats eller saknas. Dessutom kan det läggas till ljud vid tvekan, till exempel "eeh..." eller ljud kan uttalas extra länge för att talaren ska få tid att tänka efter. Allt detta innebär svårigheter för automatisk taligenkänning. [Blomberg & Elenius, 2005]

## Indirekta miljöbetingade faktorer

Enligt Gellatly (1997) finns det olika faktorer som påverkar talet. Om en användare upplever trötthet, stress eller en hög arbetsbörda kan sättet att tala förändras. Om användaren arbetar

med ett system som kräver en kort paus mellan varje yttrande kan användaren under ovanstående förutsättningar återgå till sitt naturliga flyt i talet, vilket påverkar igenkänningen negativt. Också miljön som användaren befinner sig i kan påverka variabiliteten i talet. Fysiska påfrestningar som värme, vibrationer och accelerationer ökar talets variabilitet. Likaså kan mentala påfrestningar som ångslan, mental trötthet och belastning påverka.

Taligenkänningssystem måste väljas med hänsyn till den variabilitet som kan förekomma. Talaradaptiva system kan ha hög prestanda även när användarens tal ändras under användningen av systemet. [Gellatly, 1997]

### **Direkta miljöbetingade faktorer**

Andra miljöbetingade faktorer som inte nödvändigtvis påverkar talaren direkt kan också inverka negativt på signalen till taligenkänningssystemet. Sådana faktorer kan vara omgivande ljud, akustiken talaren befinner sig i och störningar i signalöverföringen. För problemen finns direkta tekniska lösningar. Till exempel kan omgivande ljud isoleras bort och akustiken förbättras genom dämpande material. Mikrofoner som eliminerar oljud (eng. noise cancelling) kan användas för att minska påverkan av omgivande ljud. Det finns också olika signalbehandlingsmetoder för att rensa signalen från brus. [Gellatly, 1997]

Talarberoende system har en annan möjlighet att minska påverkan av omgivande ljud. Genom att låta användaren "träna" med systemet i miljön där det ska användas kommer det omgivande ljudet vara med och forma den mall som yttranden jämförs med när systemet sedan används. På så sätt blir det lättare att matcha yttranden mot systemets tillåtna vokabulär och systemets prestanda blir högre än om mallen skapats i en miljö som saknade det omgivande ljudet. [Gellatly, 1997; Baber m.fl., 1996]

### **Uppmärksamhet**

Det finns två olika teorier om hur människans uppmärksamhet fungerar som resurs. Båda teorierna behandlar uppmärksamheten som en reserv från vilken man drar resurser. Den ena teorin kallas singel-resurs teorin och behandlar uppmärksamhet som en enda reserv. Ju mer uppmärksamhet en eller flera uppgifter kräver desto mer dras från reserven. Så länge den totala efterfrågan inte överstiger tillgången påverkas inte prestationen negativt. Den andra teorin kallas multipel-resurs teorin. Den skiljer på uppgifter som använder sig av olika sorters in- och utsignaler. Teorin säger att uppgifter som inte delar samma sorts in- och utsignaler inte påverkar varandra i lika hög grad. Som exempel kan tas uppgiften att köra en bil som kombineras med uppgiften att byta radiostation. Framförandet av bilen har i huvudsak visuell insignal; synintryck av omgivningen, vägen och trafiksituationen, och manuell utsignal; styrning med ratten, blinkers, gas, broms osv. Likaså har bytet av radiostation visuell insignal; stationsnamnet visas på en display, och manuell utsignal; stationen ändras med ett vridreglage. Uppgifterna inverkar då negativt på varandra. Om radion istället skulle kunna kontrolleras med röstkommandon och ge auditiv feedback skulle inte köruppgiften påverkas lika mycket. [Gellatly, 1997]

Teorin om multipla resurser säger att det är tre saker som påverkar prestationen då det finns flera simultana uppgifter; hur information presenteras, hur informationen behandlas i hjärnan och hur svaret sker. [Gellatly, 1997]

### **Felhantering**

Det är viktigt att ett system för automatisk taligenkänning har en pålitlig metod för att rätta till fel för att användare ska känna tillförlitlighet till systemet. Vilken metod som är lämplig att

använda beror på vad det är för applikation och viken feedback systemet ger användaren. Det måste finnas en balans mellan att interaktionen med systemet kan ske flytande och besväret med att korrigera fel. Att välja rätt metod för felkorrigering för en applikation kan leda till ökad tillfredsställelse hos användaren. [Gellatly, 1997]

Gellatly (1997) ger ett flertal olika alternativ. En metod är att låta användaren säga "ja" eller "nej" efter varje ord eller fras har känts igen och presenterats för användaren. Ett "nej" skulle då ta bort ordet eller frasen och användaren få upprepa sig. Den här metoden skulle vara väldigt tidsödande och irriterande för många applikationer.

Metod nummer två är att användaren bara reagerar på feligenkänningar genom att säga t.ex. ändra eller ångra och inte behöver godkänna uttalandet som systemet känt igen korrekt. Den här metoden kan utökas till att omfatta fraser eller hela meningar. De måste då lagras i en buffert för redigering innan de skickas för att utföras av systemet. [Gellatly, 1997]

Den tredje metoden går ut på att användaren får återupprepa en fras där det förekommer fel. Det är sen upp till systemet att avgöra vilket uttalande som korrigeras. Detta kräver en viss "intelligens" av systemet. [Gellatly, 1997]

Den fjärde och sista metoden går till så att systemet identifierar otillåtna uttalanden och korrigerar dem till det mest troliga tillåtna uttalandet. Detta kan ske utan användarens inblandning eller genom att användaren ges möjlighet att välja mellan ett antal alternativ. Den här metoden används av stavningskontrollen hos ordbehandlare. [Gellatly, 1997]

## **Feedback**

I ett system som styrs av automatisk taligenkänning måste användaren få feedback eller återkoppling för att kunna rätta till igenkänningsfel. Det kan ges på två olika nivåer som kallas primär och sekundär feedback. [Gellatly, 1997]

Vid primär feedback svarar systemet direkt på ett uttalande, t.ex. hörs musiken från cd-spelaren efter att instruktion getts om att starta uppspelning eller föraren av en skogsmaskin ser aggregatet gripa tag i ett träd efter att ha gett instruktion om detta. [Gellatly, 1997]

Sekundär feedback uppstår då användarens instruktioner eller kommandon repeteras tillbaka för verifikation innan systemet utför någon handling. Som exempel kan nämnas automatiska biljettbokningssystem på telefon där användaren får verifiera systemets fråga "Vill du beställa en biljett från A till B med avgångstid 12.00?" med ett "Ja" om det är vad användaren vill eller svara "Nej" och börja om proceduren. [Gellatly, 1997]

De olika feedbacknivåerna har olika användningsområden. Primär feedback används då det inte är helt avgörande om det förekommer fel eller om det är viktigt att uppgiften genomförs snabbt. Sekundär feedback används då tillförlitligheten är av stor betydelse, antingen på grund av säkerhet eller för att det är svårt att reparera misstag.

Feedback kan ges via visuella eller auditiva kanaler. Vad som är mest lämpligt beror på vad uppgiften är och i vilken miljö användaren befinner sig. Om en användare utför en uppgift med i huvudsak visuella insignaler, som att köra en bil, kan visuell feedback påverka huvuduppgiften negativt. Därför bör i ett sådant fall auditiv eller en kombination av auditiv och visuell feedback väljas i ett sådant fall. [Gellatly, 1997]

## Metod

### **Studie i simulator**

I sammanhang som rör simulatorer menas med begreppet fidelity hur väl simulatören representerar det som den simulerar. Begreppet kan översättas till naturtrogenhet. I en studie av Löfgren m.fl. (2008) studerades Skogforsks skogsmaskinsimulators fidelity.

När en simulator används i forskningssyfte är det viktigt att utreda dess fidelity för att kunna säga något om giltigheten av de undersökningar som görs där. I en fidelitystudie gjordes därför en jämförelse mellan Skogforsks simulator och verkligheten. Det gick till så att ungefär 500 träd i ett bestånd som var moget för slutavverkning mättes till höjd, diameter och position. Den uppmätta skogen återskapades i simulatören för att kunna göra en jämförande studie av avverkningen.

Samma förare fick så avverka skogen både i verkligheten och i simulatören. Skillnaden i tidsåtgång för kran- och aggregatarbete var  $\pm 5\%$ . Den största skillnaden i tidsåtgång för avverkningen berodde på att mindre tid ägnades åt förflyttning av skördaren i simulatormiljön. Anledningen till det var dels att kranen var något starkare i simulatören så att föraren kunde arbeta på längre avstånd från maskinen och därmed göra färre förflyttningar, dels att förflyttningarna gick något snabbare på grund av mindre inverkan av underlaget.

För knapptryckningarna noterades att de flesta skillnaderna kom från kvistningen. Det beror på att i simulatören är kvistningen idealiserad men i verkligheten kan aggregatet fastna på grova kvistar. Det leder till att föraren backar aggregatet för att ta sats för att slå av kvisten och innebär fler knapptryckningar.

Spakrörelserna visade på stora likheter mellan simulatormiljön och verkligheten. Det fanns en skillnad i storleken på utslagen för kranrörelserna som berodde på att simulatorns idealiserade hydraulik inte stämde helt med verkligheten. I verkligheten uppstod vibrationer som gjorde spakrörelserna något ryckiga. I simulatören finns inga vibrationer vilket ledde till att spakrörelserna var mjukare. [Löfgren m.fl. 2008]

### **Observation**

En observation kan göras i explorativt syfte utan att ha något speciellt syfte förutom att forskaren vill utbilda sig själv. Oftast sker observationen av mänsklig prestation i den verkliga miljön t.ex. av en skördarförare när han kör skördare i skogen. [Helander, 2006]

En metod som kallas contextual inquiry (kontextuell undersökning) går ut på att forskaren observerar en människa utföra verkligt arbete, intervjua henne/honom under tiden och försöker förstå hur arbetet går till och vilka beslut som tas. [Helander, 2006]

### **Riktad öppen intervju**

Lantz (2007) delar i sin bok "Intervjumethodik" upp intervjuer i fyra olika former och grader av struktur. På den ena änden av skalan finns den strukturerade intervjun, följt av den halvstrukturerade, den riktat öppna och längst bort i andra änden den öppna intervjun.

Den riktat öppna intervjun karakteriseras av att det är en kvalitativ form, där intervjuaren definierat ett antal kvaliteter och intressanta frågeområden. Inom dessa områden känner intervjuaren empatiskt av vad respondenten upplever som mest intressant att berätta om.

Genom att undersöka respondentens subjektiva upplevelser söker intervjuaren utöka sin kunskap om fenomenets kvaliteter. [Lantz, 2007]

Intervjuformen används när syftet är att utöka förståelsen av ett fenomen. Undersökaren söker en persons subjektiva upplevelse av ett fenomen. Intervjupersonen utgår från sina egna upplevelser och tillåts att styra intervjun mot det som upplevdes som betydelsefullt. Den bild som undersökaren får är färgad av intervjupersonens uppfattning av verkligheten. På grund av att intervjupersonerna kommer att välja att fördjupa sig i olika kvaliteter kommer också intervjuerna att se olika ut. [Lantz, 2007]

Ett mindre antal intervjupersoner används än vid kvantitativa undersökningar som behöver ett stort antal respondenter för att säkerställa giltigheten. [Lantz, 2007]

### **Wizard of Oz**

Wizard of Oz-metoden används vid studier inom människa-dator-interaktion. Den tillåter personer att interagera med ett datorsystem som verkar vara självständigt, men som i själva verket helt eller delvis simuleras av en människa. Till exempel kan en försöksdeltagare fås att tro att den kommunicerar med ett datorsystem genom ett talgränssnitt, men i själva verket skrivs deltagarens ord i hemlighet in i systemet av en annan människa (the wizard – trollkarlen) som befinner sig i ett annat rum. Resultaten från en Wizard of Oz-studie kan sedan användas för utvecklingen av ett verkligt autonomt system.

Wizard of Oz-studier har funnit sitt största användningsområde för simulering av gränssnitt som använder sig av naturligt tal. [Salber & Coutaz, 1993]

Ofta rör det sig om system där användaren söker information och matar in data av något slag t.ex. almanackor, guidesystem eller biljettbokning. I fallet med ett naturligt talgränssnitt, kan resultaten av studien till exempel vara information om vilka ord och uttryck som används i den aktuella situationen, samt hur det är naturligt för försöksdeltagarna att använda tal för att kommunicera med systemet.

I praktiken går en Wizard of Oz-studie av ett naturligt tal-system till så att försökspersonen informeras om att systemet kan styras med hjälp av tal. Försökspersonen pratar med systemet genom en mikrofon och i många fall förs en dialog där systemet uttrycker sig med t.ex. skrivna svar eller talsyntes. Syftet med dialogen kan vara att be användaren att förtydliga, ge ytterligare information eller att presentera resultatet. I själva verket är det trollkarlen som i ett annat rum lyssnar och efter systemets tänkta kapacitet tolkar försökspersonens yttranden och presenterar lämpliga resultat.

Data från experimentet sparas för senare analys. Naturligtvis är en inspelning av dialogen sådan information. En videospelning kan också göras för att se försöksdeltagarens reaktioner i olika lägen och för att se hur eventuella andra kontroller används tillsammans med talgränssnittet.

### **The Wizard - Trollkarlen**

Skillnader mellan människors och datorers agerande har konsekvenser som är viktiga att förstå. Till exempel skriver människor långsamt och stavar ibland fel. Datorer, å andra sidan, är snabba och gör aldrig småfel som t.ex. stavfel. Datorer är också konsekventa till skillnad från människor som är flexibla. [Dahlbäck m.fl., 1993]

I två liknande sammanhang måste ett givet kommando från försökspersonen leda till samma respons från trolldkarlen. I en situation där det simulerade systemet ska leverera en textsträng går det därför inte att låta trolldkarlen skriva svaret i samma stund. Det skulle både gå för långsamt och kunna innehålla variationer från gång till gång. [Dahlbäck m.fl., 1993] [Salber & Coutaz, 1993]

Responstiden måste stämma överens med försökspersonens föreställning om hur lång tid det borde ta. För långsam respons kan leda till att försökspersonen undviker att använda sig av funktionen eller tror att systemet är överbelastat. För att uppnå detta måste trolldkarlen träna på sin uppgift och ges underlättande verktyg. Ett exempel på ett sådant verktyg är, i fallet då en textsträng ska levereras, en samling färdiga texter som trolldkarlen kan välja från. Samma förfrågan skulle då snabbt och fritt från stavfel generera identiska svar om och om igen. [Salber & Coutaz, 1993]

## Dialogen

Illusionen av att kommunicera med ett datorsystem, inte en annan människa, är viktig för att en studie ska ge relevanta data. Det finns flera skillnader mellan dialog människa till människa och människa till maskin som påverkar innehållet och formen av dialogen. En människa kanske inte känner något behov av att vara artig mot en dator, genom till exempel indirekta förfrågningar, på samma sätt som mot en annan människa. Det är mycket möjligt att det är effektivt att vara artig i kommunikation med andra människor men med en dator är det mer tveksamt. [Dahlbäck m.fl., 1993]

## Uppgiften

Försöksdeltagare måste på något sätt få en uppgift att genomföra under försöket. Det har visat sig att försökspersoner påverkas av hur en uppgift formuleras. Det finns flera olika hänsyn att ta. Är uppgiften för detaljerat formulerad finns det risk för att försöksdeltagaren kommer att använda sig av språket som används i uppgiftsformuleringen istället för sitt eget språk [Kelley, 1984]. Om uppgiften är för enkel kommer den inte ge tillräckligt mycket data att analysera.

Försökspersoner kan också påverkas av ordningen som information ges i uppgiftsbeskrivningen. Om en uppgift som består av flera olika steg presenteras för försökspersoner i en viss ordning kommer de att påverkas till att utföra uppgiften i den ordningen. [Brøseth, 2005]

Ett sätt att ta hänsyn till detta är att presentera uppgiften i form av ett så kallat scenario. Det ska kräva användning av systemet men det finns flera olika sätt att nå fram till möjligtvis flera olika godkända resultat. [Dahlbäck m.fl., 1993]

Lathrop m.fl. ger nedanstående exempel, fritt bearbetat och översatt, på hur en uppgift kan formuleras för olika system som kan finnas i en bil; ett navigationssystem, en musikspelare och ett system med information om restauranger.

Uppgift A: Använd navigationssystemet för att ta dig från var du befinner dig till din destination.

Uppgift B: Ta reda på vad för sorts musik som är tillgänglig i databasen, spela låten "Louie Louie" med "The Sonics" och pausa låten.

Uppgift C: Du och en vän skulle vilja äta en fin middag i den flotta staden Jackson. Ni är båda sugna på krabba . Dessutom vill det till att servitörerna visar sig från sin bästa sida eftersom din vän är något av en snobb vad det gäller service. Hitta en restaurang som passar.

I uppgift A sa deltagare ofta saker som, "Vad finns i den där bygganden, vem gjorde den där statyn och kan du få trafikljuset att slå om? Naturligtvis är det bortom systemets förmåga att svara på sådana frågor och därför är det onödigt att samla in sådana språkdata.

I uppgift B använde deltagare ofta exakt det språk som användes i uppgiftsbeskrivningen. Till exempel "Vad finns det för någon musik i databasen och kan du spela låten "Louie Louie" med "The Sonics" och pausa den.". Här använder visserligen deltagaren språk som rör sig inom rätt område men det är inte mycket poäng med att samla in språk som redan finns i uppgiftsbeskrivningen.

I uppgift C finns det inbakat fyra mål att uppnå men språket för att nå dit är inte givet. De fyra målen är att hitta en finare restaurang(1) som serverar skaldjur(2), ligger i Jackson(3) och har bra service(4). Målen är klara men språket som används för att nå dit är helt och hållet upp till varje deltagare. [Lathrop m.fl.]

## **Försökspersonerna**

Försökspersonerna ska vara representativa för de tänkta användarna av systemet.

## **Intervju och debriefing**

Efter en genomförd Wizard of Oz-simulering genomförs normalt en intervju eller enkät med försöksdeltagaren. Syftet är att samla in åsikter om det simulerade systemet. Finns det delar som utmärker sig genom att fungera bra eller dåligt och förslag till förbättringar.

Den kanske enskilt viktigaste frågan som intervjun ska ge svar på är om försöksdeltagaren misstänkte att en människa låg bakom systemet. Detta är viktigt för att kunna bedöma validiteten i insamlad data.

En Wizard of Oz-studie innebär ett bedrägeri gentemot försöksdeltagaren. Naturligtvis innebär detta vissa etiska problem. Behovet av ett bedrägeri kommer av att människor kommunicerar annorlunda med en dator än med en annan människa. Skillnaderna är också av sådan art att de sker omedvetet och det går därför inte att samla in pålitliga data om försöksdeltagaren skulle vara medveten om trollkarlens inblandning men låtsas att kommunicera med en maskin.

Efter en genomförd studie är det viktigt att göra klart för försöksdeltagaren hur försöket gått till och förklara varför. Det är också lämpligt att erbjuda att förstöra insamlad data om deltagaren skulle önska det. Det verkar dock inte vara någon större risk för att deltagare skulle ta illa upp. En anledning till detta är att de inte på något sätt har försatts i en pinsam situation. [Dahlbäck m.fl., 1993]



## **Genomförande**

### ***Fältobservation och intervju***

En observation i fält av hur avverkning med ett tvåmaskinsystem går till genomfördes under en dag i de uppländska skogarna. Under observationen gavs tillfälle att inifrån förarnas hytter uppleva hur arbetet går till. Samtidigt genomfördes en ostrukturerad intervju .

Under dagen satt författaren med i två olika skördare. En av dem var en medelstor skördare som under observationen ägnade sig åt gallring. Den andra skördaren var av största modell och slutavverkade ett bestånd.

Även två skotare besöktes. Den ena skotaren skulle få service så därför gavs bara tillfälle till ett kortare besök under avlastningen. I den andra skotaren fanns tillfälle att observera hela arbetscykeln från lastning och förflyttning i en gallrad skog till avlastning.

Samtidigt som observationerna genomfördes en intervju som dels syftade till att få ut mer information om arbetet; vilka val förarna ställs inför och beslut de fattar dels mer specifikt om reglageanvändning och arbetsmiljön som helhet.

### ***Wizard of Oz-simulering***

#### **Trollkarlens arbetsplats**

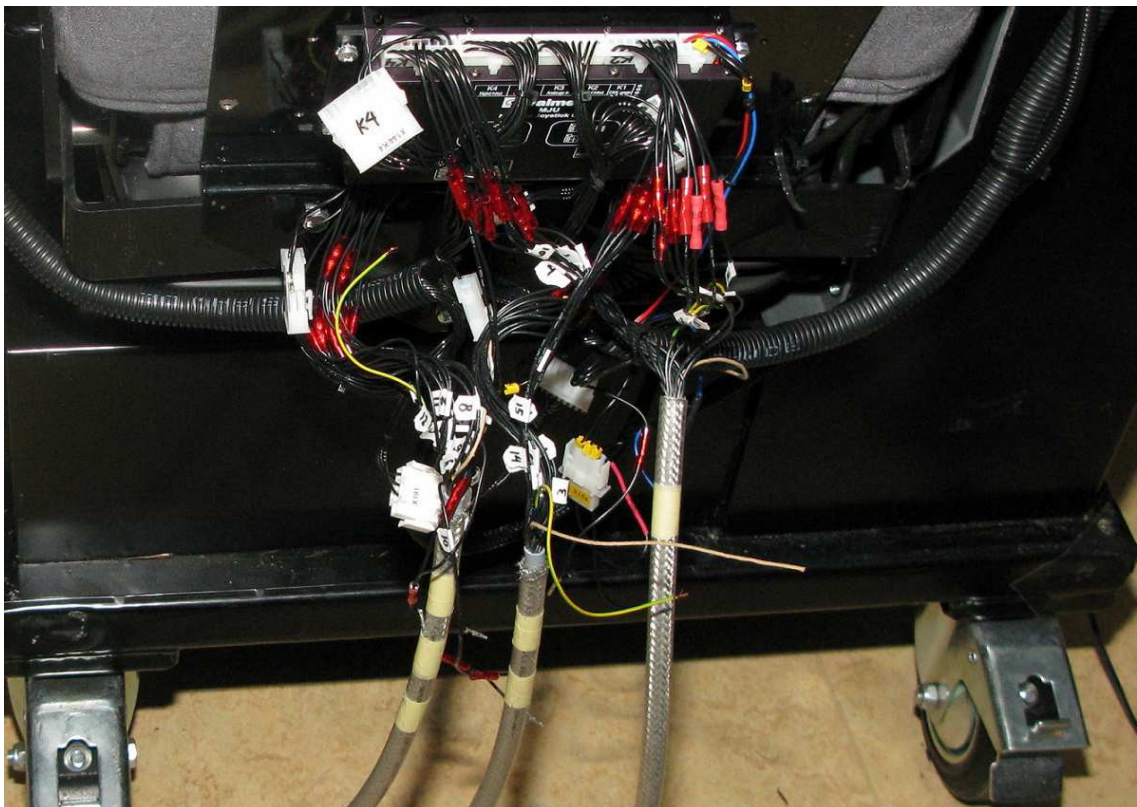
För att anpassa simulatortill att kunna användas till en Wizard of Oz-studie gjordes en del modifieringar och tillägg. En vanlig stol med armstöd försågs med likadana spakar som de som sitter på stolen i simulatort. Spakarna är fast monterade och kan alltså inte vinklas framåt, bakåt eller i sidled.

Kablar drogs från simulatorrummet till ett närliggande rum för att kunna koppla ihop trollkarlens stol med simulatort. Resultatet av kopplingarna var att alla knappar på spakarna endast kan kontrolleras från av trollkarlens stol medan styrningen av kranen och framförandet av skördaren fortfarande sköts från stolen i simulatort.

Genom att ett headset med mikrofon, som bärs av försökspersonen, kopplades till en högtalare i det närliggande rummet kunde trollkarlen hela tiden höra vad försökspersonen sade.



*Trollkarlens stol.*



*Koppling till simulatorm via skördarstolen.*

Framför sig hade trollkarlen också en skärm som visar samma apteringsdata som försökspersonen såg i simulatören. Det fanns också en duk, där den främre bilden av simulatormiljön projicerades. Eftersom hytten hela tiden roterade med kranen i den maskin som användes i simuleringen skedde all upparbetning av träden på den skärmen och det gav en bra bild av förarens arbete.

En kamera installerades i taket av simulatören framför stolen. Det var meningen att den skulle filma hur försökspersonen uppträdde under försöket. Samma ljudsignal med försökspersonens tal, som gick till trollkarlens högtalare, var också ljudspår till den inspelade filmen.

Under försöket spelades också bilden från den främre skärmen i simulatören in.

Trollkarlen befann sig i ett rum ett tiotal meter från simulatören. Knapparna på spakarna var, via tre kablar, kopplade till simulatören och ersatte knapparna på förarstolens spakar som inte längre fungerade.



*Trollkarlens arbetsplats.*

I rummet fanns två högtalare. Genom dessa fick trollkarlen sina instruktioner som försökspersonen talade in i mikrofonen. Försökspersonerna gavs ingen annan feedback annat än den direkta feedback de fick av att observera aggregatet. Detta bestämdes på förhand, eftersom det var uppenbart att det annars skulle gå alldeles för långsamt.

## Skogen

Skogen som användes under försöket bestod av 35 träd, till största delen tall och gran men också en liten del björk. För att göra uppgiften mer varierad hade träden olika skador. En tredjedel av granarna hade rotröta och varannan hade en tvärkrök. Tallarna hade i snitt en tvärkrök eller grov kvist per stam. En del träd hade ingen skada, ganska många hade en skada och en del hade två skador.

Tvärkrökar var bara tillåtna i ändarna av timmerstockarna. Förarna måste ta hänsyn till detta under apteringen för att få ut maximalt värde av trädet. Om det inte gick att planera så att kröken hamnade vid änden av en stock kommer stocken, som innehåller kröken och kunde ha använts till timmer, istället att gå till massaved. Tvärkrökarna markerades med en röd ring runt stammen på det ställe kröken satt.



*Röd och blå snitsel representerar tvärkrök respektive grov kvist.*

När ett träd har grova kvistar kan det hända att aggregatet inte lyckas slå av dem på första försöket. Föraren måste då backa aggregatet för att ta sats och försöka slå av kvistarna på nytt. I simulatormiljön yttrade sig kvistarna så att aggregatet fastnade på första försöket. Om föraren backade tillbaks och lyckades uppnå tillräckligt hög hastighet gick grenarna av på nästa försök. Grov kvist markerades med en blå ring runt stammen.

Rotröta innebär att det från roten och en bit upp har uppstått röta i trädets centrum. Normalt blir föraren uppmärksam på att trädet är utsatt för röta genom att spånen vid fällningen av trädet är bruna eller precis efter fällningen när snittet är blottat. Föraren tar hänsyn till rötan genom att kapa bort den drabbade delen som i detta scenario klassades som massaved. För att maximera värdet vill då föraren kapa bort så lite som möjligt för att få ut så mycket friskt

timmer som möjligt ur stammen. Rotröta finns i simulatormiljön och syns som en cirkulär brun fläck på snittytan av ett kapat träd.



*Rotröta.*

Det finns i verkligheten fler sorters skador som till exempel att stammen delar sig och slutar i en dubbeltopp. Eftersom det kan vara svårare för försöksdeltagaren att låtsas hantera två toppar än en krök och att dubbelkrökar är ett mer ovanligt fel finns de inte med i försöksskogen.

## **Försökspersoner**

Som försökspersoner användes sex professionella skördarförare med mångårig erfarenhet av att köra engreppsskördare. Alla förarna körde också just nu skördare från Valmet med samma typ av spakar som sitter i simulatören.

## **Inställningar av maskinen i simulatören**

Inför studien ställdes maskinen in av Anders Mörk som arbetar med utbildning av skördarförare i körteknik. Inställningarna gjordes inte efter hur de brukar se ut i verkligheten utan efter hur maskinen i simulatören faktiskt reagerar. Hur en maskin är inställd är ju naturligtvis upp till varje enskild förarens tycke och smak men målet med inställningarna var att maskinen skulle vara lättkörd, ha hög precision och mjuka rörelser.

I övrigt var maskinen inställd för så lite automation som möjligt. Till exempel var förarna tvungna att välja trädslag vid varje träd. I verkligheten ställs oftast maskinen in så att föraren bara behöver välja trädslag om det skiljer sig från närmast föregående träd. Anledningen till detta var att förarna skulle få prata så mycket som möjligt.

## **Introduktion**

Försöksdeltagaren välkomnades och gavs en kort presentation av simulatören och de fick information om tidigare studier i människa maskininteraktion.

Försöksledaren berättade för deltagaren om upplägget av studien och att ringarna på träden i beståndet representerade krökar och grova kvistar och om rötan.

## **Övning**

Försöksdeltagaren fick först köra i simulatören med konventionell styrning för att vänja sig vid hur simulatören fungerar. Det finns många möjligheter till inställningar för kran och aggregat vad gäller till exempel hastighet och acceleration. Varje förare var van vid sina egna inställningar och behövde tid till att vänja sig vid de inställningar som fanns i simulatören. Ytterligare ett syfte med övningen var att föraren skulle bli van att hantera de skador på träden som fanns i form av de röda och blå ringarna runt träden.

Det fanns ingen i förväg preciserad tid för hur lång tid som gavs för övning utan det varierade efter hur lätt de hade att anpassa sig till att köra i simulatören.

## **Presentation av röststyrningsstudien**

Efter att försöksdeltagarna vant sig vid att köra skördare i simulatören gavs information om hur försöket med röststyrningen kommer att gå till. Försöksledaren berättade att under försöket kommer kranen att styras som vanligt men att aggregatfunktionerna kommer att styras med rösten. Deltagaren fick information om den uppgift han har att lösa under försöket.

Försöksdeltagaren fick också information om att under denna del av försöket kommer ljud och bild av föraren spelas in genom mikrofonen respektive en webbkamera monterad i taket av simulatören framför föraren.

## **Uppgiften**

Uppgiften som försöksdeltagarna utförde under försöket var att avverka de 36 träd som fanns i den aktuella skogen. Föraren uppmanades att vara noggrann och ta hänsyn till virkeskvalitet och fel som fanns på stammarna samt att sortera virket i sortimentsrena högar.

Deltagarna fick inga förslag på kommandon eller uttryck utan fick veta att de själva skulle utforska systemet och prova sig fram till vad som fungerar.

## **Körning med röststyrning**

Under försöken fanns försöksledaren i simulatorrummet för att observera försökspersonerna under avverkningsarbetet och för att svara på frågor eller ge instruktioner om det behövdes. Om en förare hittat en metod som fungerade och sedan höll sig till den en tid utan förändringar uppmanades han av försöksledaren att utforska systemet ytterligare och att försöka hitta nya tillvägagångssätt.

Försöksledaren förde anteckningar över hur deltagarna agerade i olika situationer. Anteckningarna användes som underlag för frågor vid intervjun och för den efterföljande analysen.

Det fanns inte någon i förväg fastställd tid för studien. Körningen avbröts när försöksledaren bedömde att föraren inte längre utvecklade sin användning av röststyrningen och inte svarade på uppmaningar att försöka hitta nya uttryck och strategier.

## Intervju

När försöket var över intervjuades deltagarna av försöksledaren. Syftet med intervjun var att dels reda ut hur försöksdeltagarna upplevde det att styra aggregatet med rösten, dels att försöka bekräfta eller förkasta observationer som gjorts under försöket. Intervjun genomfördes som en öppen riktad intervju enligt Lantz (2007).

Intervjun spelades in efter att försöksdeltagarna gett sitt medgivande. Anledningen till inspelningen var för att senare kunna gå tillbaka för att hämta information från intervjun. Intervjuaren behöver då bara anteckna sådan information som är tänkt att användas som underlag till följdfrågor.

Som sista moment under intervjun berättade försöksledaren för deltagaren att det inte fanns något taligenkänningssystem utan att den funktionen hade simulerats av en människa. Försöksledaren förklarade hur bedrägeriet gått till och varför det var nödvändigt. Den sista frågan under intervjun var om försöksdeltagaren någon gång misstänkte att en människa låg bakom taligenkänningen.

## Resultat

### Övning

Under övningen kände flera försökspersoner av simulatorsjuka. En förare kände så stort obehag att försöket avbröts utan att han provat röststyrningen. För en annan deltagare lindrades besvären av att skärmarna på sidorna av simulatören släcktes. För de flesta räckte det att ta en paus då och då både under övningen och under körningen med röststyrning.

Förutom problemen med simulatorsjuka vände sig förarna snabbt. De största skillnaderna upplevdes vara dels svårighet att bedöma avstånd på den tvådimensionella skärmen och dels att en del funktioner satt på andra knappar än de var vana vid. Framst rörde det funktionerna byte av kvalitet och sortiment.

### Körning med röststyrning

Fem försöksdeltagare genomförde körningen med röststyrning. Alla förarna avverkade efter försöksledarens önskemål något fler än de 36 träd som fanns i den aktuella skogen. Anledningen till detta vara att försöksledaren ville ge deltagarna mer tid till att utveckla sitt arbetssätt och prova nya ord och uttryck. 36 träd var alltså enligt försöksledarens uppfattning för få. Därför startades simulationen om och försöksdeltagarna fick börja om från början då träden tagit slut. Förarna ägnade mellan 30 och 40 minuter till att avverka skog med det röststyrda systemet.

### Observation och intervju

*Wizard of Oz-simuleringens effektivitet.*

Sist i intervjun avslöjades hur simulationen gått till och anledningen till bedrägeriet förklarades. Alla försöksdeltagarna accepterade metoden och ingen tog illa upp av att ha blivit utsatt för bedrägeriet. Ingen hade heller misstänkt att det var något annat än en dator som låg bakom röststyrningen.



Vid några tillfällen gick trollkarlen händelserna i förväg och utförde ett kommando han väntat sig istället för det som försökspersonen faktiskt gav. Antalet tillfällen som detta hände var ganska få, några enstaka per förare, och det verkar inte ha påverkat illusionen av att en dator skötte röstigenkänningen.

#### | *Allmänt intryck.*

Förarna tyckte alla att det varit en positiv upplevelse. Deras svar på frågor som "Hur upplevde du det att använda rösten för att kontrollera skördaren?" var "enkelt", "spännande" och "över förväntan".

#### | *Vid första trädet.*

När förarna för första gången skulle börja använda sig av röststyrningen hittade de snabbt ord och uttryck som fungerade.

Frågan om hur de hade resonerat inför att de skulle avverka sitt första träd med röststyrning besvarades på snarlika sätt av alla förare. De kunde säga att uttrycken de använde är de som finns i systemets manual eller i apteringsdatorns menyer som sågs som de "korrekta" uttrycken. En förare uttryckte att han använde de ord han skulle använda om han skulle lära någon annan hur arbetet går till.

Bara en gång körde en förare fast på det första trädet då hans kommando inte godtogs av trollkarlen. Föraren var då i en situation med aggregatet nerfällt och stängt runt toppen av trädet. Först försökte han med "börja om". När det inte fungerade provade han "nollställ" som heller inte fungerade. Anledningen var att trollkarlen inte var säker på vad han menade. Efter några sekunders funderingar bytte han strategi och sa istället "släpp toppen" följt av "res aggregat", vilket fungerade som han tänkt.

#### | *Tankesätt.*

Alla förare hade en liknande uppfattning om hur de ville använda röststyrningen. En knapptryckning ersattes av ett röstkommando.

#### | *Utforskande.*

När en förare lyckats avverka det första trädet observerades att han använde samma metod och uttryck också för de kommande träden. Det var ändå mycket tydligt att det skedde en viss effektivisering. Uttryck som till exempel "matning bak" blev efter några träd till "back". På samma sätt kortades "aggregat upp" till bara "upp" och likadant behandlades flera andra uttryck.

På frågan om hur de utforskade systemet svarade alla att de medvetet provade att förenkla och korta ned de uttryck de börjat med. Till exempel provades om "upp" kunde ersätta "aggregat upp". Om det fungerade användes det kortare uttrycket i fortsättningen.

I ett fall ville en förare, när han gripit tag i trädet, lägga tryck på det från sidan genom att fälla ned aggregatet och använde då kommandot "fäll". Detta tolkade dock trollkarlen som att föraren ville fälla trädet och startade då en sekvens där han kapade trädet och fällde aggregatet. Vid nästa träd provade föraren om kommandot "fäll" kunde göras ännu mer användbart. Han placerade aggregatet vid stammen så "fäll" varpå trollkarlen grep tag i trädet, kapade och fällde det. Ytterligare en funktion hade alltså innefattats i samma kommando och i fortsättningen var det så föraren använde kommandot.

### *Arbetsätt.*

Förarna fick frågan om röststyrningen gjorde att de förändrade sättet de arbetade på. Svaren varierade beroende på om föraren var van vid att använda mycket manuell kontroll, dvs. manuell matning och kapning, eller hellre använde automatisk aptering. Förarna som var vana vid mycket manuella ingripanden tyckte att röststyrningen minskade möjligheten för att arbeta på det sätt som de var vana. De kände sig tvingades att använda automatisk upparbetning i stället.

En förare uppmärksammade att han kastade om ordningen på sina kommandon. Med röststyrning var det smidigast att göra trädvalet innan aggregatet sattes an mot trädet, för att sen kommandon för att gripa, kapa, fälla och mata skulle kunna komma i en snabb följd. Annars kan den knapptryckningen ske samtidigt som han griper eller kapar trädet.

### *Svårigheter.*

Flera förare upplevde att det var svårt att växla längdmodul, dvs. växla till en annan förinställd längd. Detta används då föraren vill få till en så kort rötskadad stock, eller så lång timmerstock innan en krök som möjligt.

En annan svårighet av liknande slag var att växla mellan timmer och massaved. Förarna upplevde det som att deras kommandon inte togs emot när systemet inte reagerade snabbt nog på trollkarlens knapptryckningar. Båda dessa svårigheter berodde inte på något problem med röststyrningen. Förarna visste vad de ville göra och uttryckte det på ett sätt som trollkarlen förstod, men han lyckades inte utföra kommandot på grund av att systemet inte svarade som han förväntat sig.

Vid ett kommando som t.ex. "mata" svarade trollkarlen genom att hålla inne knappen för automatisk upparbetning. Vid en krök måste föraren ofta stanna matningen för att själv välja var snittet ska vara. De flesta gjorde detta genom att säga "stopp" eller något liknande. Vid kommandon som "bakåt" eller "framåt" gjorde trollkarlen en kort tryckning på den manuella matningen. Ofta fick föraren upprepa sig för att hamna rätt. Det fanns delade meningar bland förarna hur väl det fungerade. En tyckte att det gick för fort för att hinna säga stopp på rätt ställe medan andra inte tyckte hanteringen av krökar vållade något problem.

Flera av förarna förutsåg att det skulle innebära problem i en väldigt kvistig skog. I simulatorm hade träden oftast bara ett eller två fel per träd men i verkligheten kan en sträcka på flera meter ha grova kvistar. I kombination med krökar att ta hänsyn till gör det att föraren använder sig av många tryck för manuell matning framåt och bakåt i snabb följd. Enligt förarnas erfarenheter skulle det bli svårt att utföra den matningen på samma sätt som använts i studien.

### *Fördelar.*

Flera förare påpekade att det kändes skönt att inte behöva arbeta med fingrarna.

### *Tillit.*

Alla förare kände att de kunde lita på systemet. Ingen kände någon osäkerhet över att systemet skulle fortsätta att känna igen och reagera som förväntat på ett kommando som använts tidigare. Flera förare påpekade dock det skulle vara nödvändigt att kunna stänga av röststyrningen för att ta över med knappar istället. Anledningen var dels att kunna fortsätta

avverkningen om röststyrningen inte fungerade av någon anledning och dels att förare ofta pratar i telefon samtidigt som de arbetar.

#### *Feedback.*

Ingen av förarna upplevde det som ett problem att det inte fanns någon indirekt feedback. En förare kommenterade "Nej, det behövs inte. Man har ju aggregatet framför sig."

## **Diskussion**

### ***Simulatorn***

I en studie som genomförs i en simulator är det naturligtvis relevant att fråga sig hur simulatorn påverkar studiens giltighet. Varje studie påverkas av simulatormiljön på olika sätt. Löfgren m.fl. (2008) konstaterade att det vid en tidsstudie att det beträffande prestation och tidsåtgång råder stor överensstämmelse mellan Skogforsks simulator och verkligheten.

I denna kvalitativa studie om röststyrning av skördarens aggregat var det inte prestation och tidsåtgång som undersöktes utan förarnas beteende och subjektiva upplevelse.

En faktor som utan tvekan påverkade de flesta förarna var simulatorsjuka. En av de ursprungligen sex förarna påverkades så starkt under övningsmomentet att han inte kunde slutföra studien. Av de återstående fem påverkades tre så mycket att de behövde vila från simulatorn för att kunna fortsätta utan större obehag. Antagligen kände de simulatorsjuka förarna av illamåendet en stund innan de meddelade försöksledaren att de behövde en paus.

Under körningen med röststyrning utnyttjades förarnas nyfikenhet och kreativa förmåga. Det är rimligt att anta att de påverkades negativt både av illamåendet, men även efter en paus av att veta att de befinner sig i en miljö som kommer att göra dem illamående.

Att frisk luft kan minska illamående är en erfarenhet som säkert många människor har vare sig illamåendet beror på simulatorsjuka, sjösjuka, åksjuka eller har någon annan orsak. Simulatorn befinner sig i ett ganska litet rum där det vid tillfället för studien fanns sju datorer och tre projektorer igång. Dessutom befann sig försökspersonen och försöksledaren i rummet. Kort sagt blev det varmt och kvalmigt. I framtida studier skulle det vara klokt att vidta åtgärder för att förbättra luftkvaliteten i simulatorrummet.

### ***Intervjun***

I en öppet riktad intervju kommer alla respondenter inte att svara på samma frågor. Respondenten leder själv in intervjun på det som han själv tycker är mest intressant och styr därför själv delvis vad som kommer ut av intervjun. Därför kan inte ett uttalande om att två av fem förare hade en viss åsikt tolkas som att de andra tre hade motsatt åsikt. I själva verket kanske de inte alls berörde ämnet. Om två respondenter i denna studie hade motsatta åsikter så har det redovisats.

Svaren från en studie av den här storleken, med fem försökspersoner, kan inte sägas spegla åsikter eller beteenden hos alla Sveriges skördarförare, som är den population försökspersonerna är tagna från, med någon statistisk säkerhet. Många åsikter och beteenden kanske delas av den större populationen med det kan också finnas åsikter och beteenden som är vanliga i den större populationen som inte framkommit i studien. Det som kommer ut ur

den här intervjun är att intressanta fenomen vid röststyrning av ett skördaraggregat har uppmärksammats och förarnas reflektioner över av försöksledaren tillsammans med förarna.

Författarens prestation i rollen som intervjuare var av stor betydelse för de resultat som kom ur intervjuerna. Intervjuteknik är ett hantverk som tar lång tid att bemästra och även om en viss utveckling skedde måste det sägas att författaren fortfarande är en relativt oerfaren intervjuare. Respondenterna var naturligtvis olika. En del tog större eget initiativ och var mer drivande i intervjun än andra. Gemensamt för alla var att de upplevdes som intresserade av ämnet och gärna delade med sig av sina upplevelser och åsikter.

### **Wizard of Oz-metoden**

Att Wizard of Oz simuleringen av röststyrningen fungerade, d.v.s. trollkarlen kunde utföra passande kommandon och ingen misstänkte att en människa låg bakom, är ett intressant resultat i sig. De flesta Wizard of Oz-studier utförs i sammanhang där användaren gör ett relativt litet antal yttranden och trollkarlen har relativt gott om tid på sig att leverera ett svar. Användaren är generellt sett ingen expertanvändare och är inte insatt i hur systemet fungerar.

I den här studien var användarna professionella skördarförare och väl insatta i hur skördaren fungerar. Dessutom levererade de kommandon i en snabb följd som krävde en omedelbar reaktion från trollkarlen.

Trollkarlen var i det här fallet en före detta skördarförare som dessutom hade stor erfarenhet av att köra simulatoren. Han var därför väl bekant både med spakarnas knappar och med hur avverkningsarbetet går till. Detta gjorde att han kunde förbereda sig för kommandon som han förväntade sig skulle komma. Därmed kunde han reagera snabbt men det bidrog också till att trollkarlen ibland, när föraren sa något oväntat, tryckte på fel knapp. Det stora flertalet av trollkarlens svar var dock korrekta och snabbheten bidrog säkert till försökspersonernas positiva inställning till att använda sig av röststyrningen.

De svårigheter som uppstod med att ändra modul och att växla mellan timmer och massaved kunde ha undvikits genom att ha gjort fler provkörningar av Wizard of Oz-simuleringen för att ge trollkarlen mer övning och rätt ut alla problem innan studien. Alternativt skulle fler förare kunnat användas och gett tid till att reda ut problemen efter hand.

### **Inspelad data**

Videon av förarna, när de satt i simulatoren, visade sig vara till liten nytta vid den efterföljande analysen. Däremot var ljudinspelningarna av förarnas kommandon användbara.

Särskilt användbart var kombinationen av ljudinspelning och inspelningen av den främre bilden i simulatoren. Då kunde försökspersonernas yttranden sättas i samband med vad som hände. Tyvärr gjordes bara en sådan inspelning under de två sista försöken på grund av tekniska problem.

### **Framtida forskning**

Det finns en begränsning för hur mycket det går att lära sig av Wizard of Oz-studier. För att verkligen studera hur förarens prestanda påverkas av röststyrning av aggregatet och hur arbetsbelastningen påverkas måste ett fullt fungerande system utvecklas och utvärderas i fält.

Det skulle dock vara möjligt att genomföra en Wizard of Oz-studie som ger en god indikation på skillnaden i förarens arbetsbelastning och prestanda vid användning av röststyrning jämfört med konventionell styrning.

I en sådan studie skulle försöksdeltagarna behöva längre tid på sig att öva med röststyrning. Trollkarlen upplevde redan under den här studien att den mentala påfrestningen var hög. Att låta försökspersonerna öva en längre tid, möjligtvis under flera olika tillfällen, skulle innebära större svårighet för trollkarlen att vara konsekvent och uppehålla illusionen av ett fungerande röststyrningssystem. Möjligheten att spela in förarens tal tillsammans med den främre bilden från simulatortorn som användes i den här studien skulle dock kunna användas för att studera tidigare övningsstillfällen och underlätta trollkarlens arbete. Mätdata skulle kunna vara t.ex. upplevd arbetsbelastning och tidsåtgång per avverkat träd.

För kvantitativa studier av röststyrning i simulatortorn kan den skillnad i antal knapptryckningar som noterades vid kvistningen i fidelitystudien, Löfgren m.fl. (2008), vara ett problem. De idealiserade stammarna i simulatormiljön skulle antagligen leda till resultat som gynnade röststyrningen. Särskilt då flera försökspersoner i den här studien påpekade att de trodde att röststyrning skulle fungera sämre i mer svårarbetad skog skulle det vara en intressant situation att testa röststyrningen i.

Ett annat intressant problem med studier av röststyrning är att det är svårt att avgöra hur det ska användas på bästa sätt. I den här studien ersattes alla knappar av röststyrning. Det är troligt att det finns mer effektiva sätt. Kanske skulle en kombination av konventionell styrning och röststyrning vara bättre. Hur den kombinationen ser ut skulle antagligen variera beroende på förare och på egenskaperna hos det bestånd som avverkas.

Det skulle kunna vara möjligt att genomföra en Wizard of Oz-studie där föraren ges möjlighet att själv, för varje handling, avgöra om han vill använda sig av sin röst eller en knapptryckning.

Att starta en utveckling av ett fungerande röststyrningssystem är ett stort och kostsamt arbete i förhållande till de förberedande studierna som är möjliga att göra med hjälp av Wizard of Oz-metoden. Det är därför rimligt att genom sådana studier försöka samla så mycket kunskap som möjligt för att avgöra om röststyrning är en lämplig metod för kontroll av engreppsskördare.

## **Slutsatser**

En skördarförare som ges fria händer att använda röststyrning för att styra aggregatets funktioner kommer antagligen att ersätta varje enskild knapptryckning med ett sedan tidigare välkänt uttryck som beskriver knappens funktion.

Metoden går snabbt att lära sig och det är tydligt att det snabbt sker en medveten effektivisering genom att göra uttrycken kortare. Uttryck som innehåller två ord oftast ner till ett, till exempel blir ”aggregat upp” till ”upp”.

Försökspersonerna var positivt inställda till röststyrningen, men förutsåg att det skulle passa mindre bra i svårarbetad skog som kräver mycket manuell aptering. Röststyrning gynnar därmed användning av automatiska sekvenser.

I studien ersatte röststyrning knapparna fullt ut och ingen kombination med knappar eller någon annan metod för styrning av aggregatet provades. Det är mycket möjligt att det finns en sådan kombination, där röststyrningens förtjänster utnyttjas bättre.

Wizard of Oz-simulering av röststyrning av aggregatet på en engreppsskördare fungerar och är ett effektivt verktyg för att skaffa mer kunskap om ämnet.

Röststyrning är en användbar teknik med potential att förbättra arbetsmiljön för maskinförarna. Fler studier behöver dock genomföras för att undersöka hur arbetsbelastning och prestation påverkas, samt för att förfina användningen av Wizard of Oz metodiken för skogsmaskinapplikationer.

## Referenser

Attebrant, Monica. Mathiassen, Svend Erik & Winkel, Jörgen. "Belastningsergonomi och rationaliseringar" i "Konsensusrapporter rörande kunskapsläget om arbetsmiljön i skogsmaskiner" (1998)

Blombeg, Mats & Elenius, Kjell (2005) "Automatisk igenkänning av tal", Institutionen för tal, musik och hörsel, KTH.

Baber, Chris & Noyes, Jan (1996) "Automatic speech recognition in adverse environments." i Human Factors. Volume 38. Issue 1.

Brøseth, Heidi. (2005) "The influence of written task descriptions in Wizard of Oz experiments" Proceedings of the 15th NODALIDA conference, Joensuu 2005

Dahläck, Nils. Jönson, Arne. Ahrenberg, Lars. (1993) "Wizard of Oz studies – Why and how." Natural language processing laboratory. Department of Computer and information Science. Linköping University.

Eklund, Jörgen & Cederqvist, Tony. "Belastningsergonomi" i "Konsensusrapporter rörande kunskapsläget om arbetsmiljön i skogsmaskiner" (1998)

Gellatly, Andrew William (1997) "The use of speech recognition technology in automotive applications" Blacksburg, Virginia.

Helander, Martin (2006) "A Guide to human Factors and Ergonomics" second edition.

Hägg, Göran M. (2001) "Handintensivt arbete – En belastningsergonomisk kunskapsöversikt gällande människans kapacitet och interaktion med verktyg och arbetsuppgifter." Arbete och hälsa 2001:9

Kelley, J. F. (1984) "An Iterative Design Methodology for User-Friendly Natural Language Office Information Applications"

Lantz, Annika. (2007) "Intervjumetodik" Studentlitteratur

Lathrop, Brian. Hua, Cheng. Weng, Fuliang. Mishra, Rohit. Joyce, Chen. Bratt, Harry. Cavedon, Lawrence. Bergmann, Carsten. Hand-Bender, Tess. Pon-Barry, Heather. Bei, Benjamin. Raya, Madhuri. Shriberg, Liz. (2004-) "A Wizard of Oz framework for collecting spoken human-computer dialogs: An experiment procedure for the design and testing of natural language in-vehicle technology systems.

Löfgren, Björn. Nordén, Berndt & Lundström, Hagos (2008) Arbetsrapport nr. 652 från Skogforsk

Salber, Daniel & Coutaz, Joëlle. (1993) "Applying the Wizard of Oz Technique to the Study of Multimodal Systems"

Internetreferens:  
[www.skogforsk.se](http://www.skogforsk.se)