



**Röststyrning i fabriksmiljöer:**

*En behovsanalys av röstbaserade system inom fabriksmiljöer*

**Voice control in factory environments:**

*A needs analysis for speech-based systems in factory environments*

Examensarbete inom *User Experience Design*  
Grundnivå 30 Högskolepoäng  
Vårterminen 2019

Joakim Wikberg

Handledare: Erik Lagerstedt  
Examinator: Henrik Svensson

## Förord

Jag vill bara säga *stort tack* till min handledare Erik Lagerstedt för all hjälp och feedback under arbetets gång. De snabba svaren via e-mail och intressanta diskussionerna har verkligen uppskattats. Jag vill även tacka min examinator Henrik Svensson för all den användbara återkopplingen på arbetet vid examinationstillfällena.

Mac Svan på Volvo i Skövde har också förtjänat ett stort tack för all hjälp och alla goda insikter under genomförandet av arbetet. Jag vill även tacka själva organisationen (*Volvo Group*) för möjligheten att genomföra detta examensarbete hos er och för tillgången till fabriken och intervjudeltagare.

Slutligen, tack till alla som på ett eller annat sätt hjälpt mig med detta arbete. Särskilt tack till min klasskamrat Marcus Göransson för sällskapet på Volvo och för möjligheten att bolla idéer.

//Joakim

## Sammanfattning

Digitala röstbaserade system uppkom först på 1970-talet och har sedan dess utvecklats och blivit ett allt mer vardagligt fenomen. Det finns till exempel diverse röstbaserade digitala assistenter i de flesta av dagens moderna smartphones, såsom Apples *Siri*, *Google Assistant* och Microsofts *Cortana*. Utöver konsumentelektronik, finns det flera potentiella användningsområden för denna typ av system. I detta arbete har potentialen och behoven för röstbaserade system inom fabriksmiljöer undersökts. Termen *behov* syftar i detta sammanhang på vad användare av exempelvis ett system behöver för att uppnå sina mål (till exempel vid utförandet av någon specifik uppgift). Genom att ha användarnas behov i åtanke vid utvecklingen av digitala system går det att utforma system som i större utsträckning kan tillgodose användarnas behov och hjälpa dem nå sina mål, vilket generellt sett bidrar till en positivare *användarupplevelse* (eng. *User Experience*). Behov manifesterar sig ofta som problem och frustrationer som användare kan uppleva, eller som begränsningar och möjligheter inom kontexten. Till exempel, inom detta sammanhang, skulle ett behov kunna ses som en möjlighet att förbättra eller underlätta utförandet av arbetsuppgifter med hjälp av ett röstbaserat system inom fabriksmiljöer. I detta arbete har behov undersökts empiriskt genom observationer och intervjuer med arbetare inom en fabriksmiljö. Baserat på resultaten från undersökningen kunde ett antal möjligheter och begränsningar för röstbaserade system inom fabriksmiljöer identifieras: möjligheten för arbetare att programmera/konfigurera bearbetningsenheter via röstkommandon, föra anteckningar relaterade till sitt arbete via diktering samt möjligheten att ta emot återkoppling/information via röst. Den huvudsakliga eventuella begränsningen som identifierades var ljudnivån inom fabriksmiljöer samt tekniska och praktiska utmaningar relaterade till dessa system. Baserat på de identifierade begränsningarna och utmaningarna föreslås det (bland annat) som fortsatt arbete inom ämnesområdet att dessa undersöks närmare och mer djupgående.

## Populärvetenskaplig sammanfattning

De allra flesta har säkerligen sett en eller två filmer och tv-serier med någon form av datorsystem eller artificiell intelligens som både kan förstå och tala med människor, samt utföra diverse uppgifter baserat på röstkommandon. Denna typ av system som ses i filmer och tv-serier anses ofta vara mer eller mindre *futuristisk*, och i verkligheten är nog dagens teknologi inte riktigt där än. Däremot är nog många även bekanta med de redan existerande ”smarta assistenterna” som finns i moderna smartphones. Dessa assistenter, som det går att ge röstbaserade kommandon till eller föra en konversation med, är inte science fiction, och de blir allt mer populära.

Röstbaserad kommunikation är en fördelaktig och naturlig kommunikationsform för oss människor, och med dagens teknologiska utveckling är det därmed intressant att undersöka vilka fler potentiella användningsområden som det kan tänkas finnas för dessa system. I detta examensarbete har *behov* för röstbaserade system inom fabriksmiljöer undersökts. Behov kan definieras som vad användare *behöver* för att uppnå sina *mål* inom en specifik kontext. Användarnas mål kan i sig definieras som vad användarna ämnar uppnå vid utförandet av en specifik uppgift. Vidare manifesterar sig behov ofta som möjligheter och begränsningar i kontexten relaterade till användarnas mål, alternativt som möjligheter att förbättra hur dessa mål kan uppnås. Till exempel skulle ett behov för röstbaserade system kunna ses som en möjlighet för att användare att på ett bättre eller enklare sätt kunna utföra en uppgift och uppnå sina mål med hjälp av denna typ av system. Genom att undersöka användarnas behov och ha dem i åtanke under utvecklingsprocessen går det att skapa system som i större utsträckning tillgodoser användarnas behov. Som i sin tur bidrar till att de kan uppnå sina mål, vilket är en faktor som generellt sett bidrar till en bättre *användarupplevelse* (eng. *User Experience*).

I detta examensarbete har behov för röstbaserade system undersökts empiriskt genom utförandet av en *behovsanalys*. Undersökningen genomfördes i en av *Volvo Group Truck Operations* fabrikslokaler i Skövde. Datainsamlingsdelen av undersökningen bestod av två metoder: *direktobservationer* och *semi-strukturerade intervjuer*. Data samlades in om kontexten, det vill säga fabriksmiljön, arbetarna och deras arbetsuppgifter, nuvarande användarupplevelse, tidigare erfarenhet och attityd till röstbaserade system samt vilka möjligheter/begränsningar de själva kunde se för användningen av dessa system i sitt eget arbete inom fabriken.

I litteraturen nämns det vissa utmaningar relaterade till empiriska undersökningar inom fabriksmiljöer. Dessa utmaningar märktes även vid genomförandet av detta arbete. De empiriska undersökningarna fick till exempel inte störa det pågående arbetet inom fabriken, vilket medförde vissa metodologiska utmaningar. Författaren av denna rapport fick därmed anpassa datainsamlingsmetoderna i viss utsträckning. Detta var en metodologisk kompromiss som författaren inser inte var optimal för genomförandet och undersökningen, men som samtidigt var nödvändig.

Insamlad data analyserades därefter *deduktivt* och *induktivt* med hjälp utav *affinitetsdiagram*. Resultatet består av:

- Miljöbeskrivningar av den aktuella fabriken som undersökts.
- Observerbara möjligheter och begränsningar för röstbaserade system.
- Respondenternas arbetsuppgifter samt dator- och systemanvändning.
- Respondenternas användarupplevelse relaterad till deras arbete.
- Respondenternas tidigare erfarenhet av och attityd till att använda röstbaserade system i sitt arbete, samt möjligheter och begränsningar de själva kunde se associerade med den potentiella användningen av röstbaserade system i sitt arbete.

Baserat på resultatet från undersökningen kunde följande slutsatser dras kring behov som identifieras i formen av möjligheter och begränsningar: möjligheten för operatörer att programmera/konfigurera bearbetningsenheter via röstkommandon, möjligheten att föra anteckningar såsom arbets- och felbeskrivningar via diktering samt möjligheten att få återkoppling/information via röst. Den huvudsakliga begränsningen berör ljudnivån (bullernivån) inom fabriksmiljöer. Slutsatsen som går att dra från detta är att ljudnivån allra först måste testas för att säkerställa att det ens är möjligt att använda röststyrning inom fabriken av intresse innan utvecklingen av ett system påbörjas. En annan begränsning har med huruvida det är tekniskt möjligt och praktiskt att faktiskt använda röstbaserade system för specifika uppgifter. Attityder och tidigare erfarenheter är ytterligare en intressant aspekt som kan medföra vissa begränsningar. Om användare till exempel har negativa erfarenheter eller en allmänt negativ attityd till att använda denna typ av system kan det medföra vissa begränsningar. Resultatet från undersökningen i detta arbete indikerade att respondenternas tidigare erfarenhet och attityd till röstbaserade system var blandad, men övervägande positiv.

Förslagsvis skulle nästa steg för fortsatt undersökning inom detta ämnesområde vara att testa hur effektiv tolkningen av talinmatning är i relation till bullernivåer inom fabriksmiljöer. Det hade även varit intressant att undersöka potentiella behov mer i detalj inom ett avgränsat användningsområde, till exempel avgränsat till ett fåtal arbetsuppgifter och systeminteraktioner. Alternativt skulle tekniska och praktiska utmaningar associerade med denna typ av system och vilka begränsningar de kan medföra undersökas vidare.

# Innehållsförteckning

<b>1. INLEDNING .....</b>	<b>1</b>
<b>2. BAKGRUND OCH PROBLEMPRECISERING .....</b>	<b>3</b>
2.1. Terminologin använd i rapporten.....	3
2.2. Röstbaserade systems uppkomst, potential och utmaningar.....	4
2.3. Röstbaserade system i fabriksmiljöer och andra kontexter .....	5
2.4. User Experience Design och användarupplevelse .....	7
2.5. Användarcentrerad design, användarinvolvering och behov .....	8
2.6. Användarupplevelsen i fabriksmiljöer .....	9
2.7. Problemprecisering .....	11
<b>3. METODOLOGISKA VAL .....</b>	<b>13</b>
<b>3.1. Val av datainsamlingsmetoder .....</b>	<b>13</b>
3.1.1. Direktobservation .....	14
3.1.2. Semi-strukturerad intervju.....	14
<b>3.2. Val av analysmetoder .....</b>	<b>16</b>
3.2.1. Deduktiv och induktiv analys.....	16
3.2.2. Affinitetsdiagram.....	16
<b>4. GENOMFÖRANDE .....</b>	<b>17</b>
<b>4.1. Direktobservationer .....</b>	<b>17</b>
<b>4.2. Semi-strukturerade intervjuer .....</b>	<b>18</b>
4.2.1. Utformning av intervjufrågor och manus .....	18
4.2.2. Pilottest .....	19
4.2.3. Genomförande av intervjuer.....	20
<b>4.3. Etiska aspekter relaterade till insamling och hantering av data.....</b>	<b>21</b>
<b>4.4. Analys av observationsdata .....</b>	<b>23</b>
<b>4.5. Analys av intervjudata .....</b>	<b>23</b>

<b>5. RESULTAT .....</b>	<b>25</b>
<b>5.1. Resultat från direktobservationerna .....</b>	<b>25</b>
5.1.1. <i>Rumsliga observationer</i> .....	25
5.1.2. <i>Observationer kring aktörer</i> .....	25
5.1.3. <i>Observationer kring användningen av system</i> .....	26
5.1.4. <i>Observerade potentiella möjligheter för röststyrning</i> .....	26
5.1.5. <i>Observerade eventuella begränsningar för röststyrning</i> .....	26
<b>5.2. Resultat från de semi-strukturerade intervjuerna .....</b>	<b>27</b>
5.2.1. Respondenternas arbetsuppgifter samt dator- och systemanvändning i fabriken .....	27
5.2.2. Respondenternas nuvarande upplevelse vid dator- och systemanvändning .....	27
5.2.3. Respondenternas nuvarande upplevda arbetsbelastning .....	28
5.2.4. Respondenternas nuvarande upplevelse av stress .....	29
5.2.5. Respondenternas tidigare erfarenheter av röststyrning .....	29
5.2.6. Respondenternas attityd till potentiell användning av röststyrning i sitt arbete .....	30
5.2.7. Möjligheter för röststyrning respondenterna kunde se .....	31
5.2.8. Begränsningar för röststyrning respondenterna kunde se .....	32
5.2.9. Övriga resultat från intervjuerna .....	32
<b>6. SLUTSATSER OCH DISKUSSION .....</b>	<b>33</b>
<b>6.1. Identifierade behov i formen av möjligheter och begränsningar för röstbaserade system .....</b>	<b>33</b>
<b>6.2. Förslag på fortsatt arbete inom ämnesområdet .....</b>	<b>36</b>
<b>6.3. Etiska aspekter relaterade till ämnesområdet .....</b>	<b>36</b>
<b>6.4. Arbetets bidrag till UXD-fältet .....</b>	<b>37</b>
<b>6.5. Utmaningar och lärdomar .....</b>	<b>37</b>
<b>6.6. Sammanfattning av slutsatser .....</b>	<b>39</b>
<b>REFERENSER .....</b>	<b>40</b>
<b>BILAGOR .....</b>	<b>42</b>

# 1. Inledning

Vi människor har utvecklats till att väldigt effektivt kunna uppfatta röster och tolka mening i vad som sägs (Nass & Brave, 2006). Kommunikation via tal är dessutom fördelaktigt eftersom det är en naturlig form av kommunikation för oss människor, vilket medför en lägre grad kognitiv ansträngning (Pires, 2005). Digitala *röstbaserade system* började utvecklas först på 1970-talet (Koester, 2001) och i dagsläget är de flesta av oss bekanta med de vanligt förekommande röstassistenter i dagens smartphones. Det finns dock flertalet domäner utöver konsumentelektronik där potentialen för röstbaserade system kan vara av intresse att undersöka. I detta examensarbete har potentialen och *behoven* för röstbaserade system inom fabriksmiljöer undersökts genom utförandet av en *behovsanalys*. I relation till detta har även kontexten och användarnas nuvarande upplevelse inom fabriksmiljöer undersökts. Termen *behov* syftar vanligtvis på vad användare behöver för att uppnå sina mål, och när det kommer till att undersöka dessa menar Kujala (2006) att behov ofta manifesterar som problem som hindrar eller försvårar för användare att uppnå sina mål i nuläget. Behov kan också manifestera sig som möjligheter för att *förbättra* hur användare kan uppnå sina mål - till exempel, i relation till röstbaserade system, kanske det kan vara möjligt för användarna att utföra en arbetsuppgift med större effektivitet eller tillfredställelse med hjälp av röststyrning.

“Thus we argue that user needs manifest themselves as either problems that hinder users in achieving their goals, or as opportunities to improve the likelihood of users achieving their goals.” (Kujala, 2006, s. 464).

I denna rapport definieras *kontexten* som en specifik *miljö*, *användarna* och *uppgifterna* som utförs inom den specifika miljön (SS-EN ISO 9241-210:2010). Kontexten kan bidra med insikter kring både möjligheter och begränsningar. Till exempel sätter kontexten begränsningar på vad som är realistiskt och praktiskt möjligt, och det är därmed viktigt att undersöka kontexten och även ha den i åtanke vid genomförandet av behovsanalys (Kujala, 2006). Den specifika fabriksmiljön som undersökts i detta arbete är en av *Volvo Group Truck Operations* (hädanefter refererad till enbart som *Volvo*) fabrikslokaler i Skövde. I nuläget används inte röststyrning i någon utsträckning i Volvos fabriker i Skövde, men handledaren på Volvo indikerade att det kan vara intressant att undersöka.

*Användarupplevelsen* eller *User Experience (UX)* är en term som används för att beskriva användares upplevelse vid interaktionen eller användningen av en produkt, ett system eller en tjänst. Det handlar till stor del om huruvida användaren upplevt interaktionen som positiv eller negativ. Det finns flertalet faktorer som påverkar användarupplevelsen, men några av de främsta anses vara användbarhet, funktionalitet, utseende och emotionella aspekter relaterade till interaktionen (Preece, Rogers & Sharp, 2015/2016). I detta arbete har användarnas (fabriksarbetarnas) nuvarande användarupplevelse valts att undersökas eftersom insikt i deras nuvarande upplevelse kan bidra med information om eventuella frustrationer eller problem som de upplever, vilket, som nämns ovan, är en potentiell källa när det kommer till att utröna behov.



Inom UX-området läggs det ofta fokus på att arbeta *användarcentrerat*, vilket innebär att användarnas behov och mål sätts i fokus under design- och utvecklingsprocessen. Enligt ISO-standarden för användarcentrerad design (SS-EN ISO 9241-210:2010) bidrar ett användarcentrerat tillvägagångsätt i utvecklingsprocessen till produkter som i större utsträckning skapar en positiv användarupplevelse. Det talas också ofta i litteraturen om *användarinvolvering*, vilket går ut på att användare deltar i varierande grad inom projekt. Det är vanligt att användare antar en informativ roll, vilket innebär att användare delar med sig av diverse information, till exempel åsikter, behov, med mera (Kujala, 2006). Detta arbete utgick från ett användarcentrerat tillvägagångsätt där användare (intervjurespondenter) antog en informativ roll vid datainsamlingen och undersökningen.

Vidare kan detta examensarbete ses som ett första steg i en eventuell utvecklingsprocess, ämnat att undersöka behov för röstbaserade system. Arbetet ämnar ge en generell överblick över möjligheter och begränsningar för röstbaserade system inom fabriksmiljöer. Resultatet skulle kunna användas som en grund och guida mer djupgående utforskning och analys av ämnesområdet. Resultatet av vidare undersökning och analys skulle sedan kunna användas vid en eventuell utveckling av röstbaserade system för fabriksmiljöer och därmed bidra till att dessa system i större utsträckning tillgodoser användarnas behov, vilket skulle bidra till en positivare användarupplevelse för de framtida användarna.

## 2. Bakgrund och problemprecisering

I bakgrunden redogörs det först för olika termer som används för att benämna röststyrning/röstbaserade system samt hur de används i rapporten. Därefter redogörs det för röstbaserade systems utveckling, potential, utmaningar, samt tidigare studier kring röstbaserade system. I nästa del av rapporten redogörs det för termerna *User Experience Design (UXD)* och *användarupplevelse*, samt *användarcentrerad design* och *användarinvolvering*, som är två centrala ”tillvägagångsätt”/utgångspunkter inom UXD-fältet. Därefter presenteras tidigare studier kring användares (fabriksarbetares) användarupplevelse inom fabriksmiljöer. Till sist presenteras en problemprecisering bestående av en beskrivning av syftet med detta arbete, frågeställningen och en motivering till varför detta är intressant att undersöka.

### 2.1. Terminologin använd i rapporten

I litteraturen används det flertalet olika termer för att benämna röststyrning/röstbaserade system: ”speech recognition systems”, ”voice control systems”, ”voice user interface”, ”speech interface”, med flera. Ofta syftar alla benämningar på samma sak: röstbaserade system som kan ta emot, tolka och agera på talinmatning, samt ge återkoppling via tal.

Termen ”röststyrning” är egentligen inte det bästa ordvalet, i och med att termen insinuerar att det enbart rör sig om system som *stys* via tal- och röstkommandon. Dock används, för enkelhetens skull, termerna ”röststyrning” och ”röstbaserade system” (och eventuellt andra liknande termer), synonymt i denna rapport och syftar alla på den ovannämnda definitionen av röstbaserade system.

Termen *behov* definieras som vad användare behöver för att uppnå ett eller flera mål inom en specifik kontext. Behov kan manifesteras sig på flertalet sätt, ofta som problem och frustrationer som användare upplever i nuläget och som eventuellt hindrar dem från att uppnå sina mål. Om en användare till exempel upplever någon form av problematik vid inmatning av data i ett system skulle ett behov exempelvis kunna vara att inmatningen av data i systemet ska förenklas. Kontexten kan samtidigt medföra både möjligheter och begränsningar för vad som är möjligt och det är därmed viktigt att även undersöka den och ha eventuella begränsningar i åtanke (Kujala, 2006). Termen *behovsanalys* syftar på processen att undersöka, utröna och analysera behov med hjälp av diverse metoder och tillvägagångsätt (Wikberg Nilsson, Ericson & Törlind, 2015). Denna process innefattar ofta insamling av information om användarna, aktiviteter och uppgifter de utför, kontexten de befinner sig i, med mera. Ur denna information kan sedan (som nämns ovan) behov utrönas från exempelvis identifierade problem och frustrationer som användare upplevt, eller som möjligheter och begränsningar i kontexten.

## 2.2. Röstbaserade systems uppkomst, potential och utmaningar

Inom alla världens kulturer är tal den primära kommunikationsformen när det kommer till en rad olika sociala interaktioner, såsom att informera, övertala och bygga relationer med varandra (Nass & Brave, 2005). Vi människor är experter på att bearbeta tal och utrona olika personers röster. Till exempel kan foster redan innan födseln urskilja sin moders röst från andra individers. Vi människor har alltså genom evolutionens gång utvecklats till att väldigt effektivt kunna uppfatta, bearbeta och tolka tal och röster, vilket gör det till en fördelaktig kommunikation- och interaktionsform (Nass & Brave, 2005).

Hur ser det då ut i nuläget när det kommer till att kombinera teknologi och tal? Röstbaserade system började utvecklas redan på 1970-talet och under slutet av 80-talet hade systemen utvecklats till att kunna uppfatta och tolka flertalet ord och fraser i följd. Dock var användare tvungna att säga dem långsamt och tydligt, med pauser mellan varje ord för att systemen skulle uppfatta dem korrekt. Under 90-talet utvecklades dessa system ytterligare och år 1997 kom det första *Continuous Speech Recognition System* som lät användare använda sin naturliga röst och talhastighet vid interaktionen (Koester, 2001). Dock uppstod det ofta problem vid användandet av dessa äldre talgränssnitt; det var inte helt ovanligt att de uppfattade ord och fraser fel (Preece, Rogers & Sharp, 2015/2016).

I dagsläget har röststyrda system och gränssnitt utvecklats betydligt och blivit mer sofistikerade. Populariteten har dessutom ökat i och med uppkomsten av ”smarta assistenter” i smartphones, till exempel Apples *Siri*<sup>1</sup>, *Google Assistant*<sup>2</sup> eller Microsofts *Cortana*<sup>3</sup>. Däremot finns det en lång väg att gå innan dessa smarta assistenter kan tolka tal lika väl som vi människor kan (Preece, Rogers & Sharp (2015/2016). Rogowski (2012) nämner även att röststyrda system troligtvis aldrig kommer utvecklas till att ha samma känslighet och reliabilitet som det mänskliga örat när det kommer till att uppfatta ord och meningar. Rogowski (2012) menar att det alltid kommer finnas en risk att dessa system missuppfattar vad som sägs. Vi människor uppfattar däremot inte heller alltid alla fonem och ord korrekt, men baserat på en semantisk förståelse av kontexten kan vi ofta ”gissa” oss till vad som sägs, vilket Rogowski (2012) menar är en utmaning för röstbaserade system. Vidare nämner Rogowski (2012) följande utmaningar när det kommer till taligenkänningsystems tolkning av talinmatning:

- Pauser i tal, som kan bero på att talaren tvekar på det hen tänker säga.
- Upprepningar av ord och andra fraser som inte bidrar med någon extra mening till meningar.
- Utelämnade ord, eller ord som talas otydligt som är viktiga för meningens mening.

---

<sup>1</sup> Apples *Siri* (<https://www.apple.com/siri/>) [2019-05-03]

<sup>2</sup> *Google Assistant* (<https://assistant.google.com/>) [2019-05-03]

<sup>3</sup> Microsofts *Cortana* (<https://www.microsoft.com/en-us/cortana>) [2019-05-03]

Det finns tillfällen när röstbaserade system kan vara särskilt fördelaktiga; Nielsen (2003) nämner exempelvis att röstbaserade system är bättre lämpade än grafiska gränssnitt när det kommer till situationer där användarnas händer eller blick är otillgängliga för interaktionen, vilket till exempel skulle kunna vara fallet i en fabriksmiljö där användare exempelvis kan vara upptagna med andra uppgifter och inte har möjlighet att använda händerna till att interagera via exempelvis mus eller tangentbord.

Fortsättningsvis, redogör Ayres och Nolan (2006) för tre huvudsakliga användningsområden som röststyrning kan kategoriseras i: *kommando och kontroll* (eng. *command and control*), *diktering* (eng. *dictation*) och *autentisering* (eng. *authentication*). När det kommer till kommando och kontroll rör det sig om system som möjliggör för användare att kontrollera enheter eller objekt i sin omgivning genom röstkommandon. Diktering handlar om att system kan transkribera det användaren säger. Autentisering berör möjligheten för identifiering av användare via röst, vilket görs möjligt genom att varje individs röst är unik. Vidare kan taligenkänningssystem klassificeras som antingen *beroende* eller *oberoende* av talaren. Om systemet är beroende av talaren innebär det att systemet måste tränas i att känna igen användarens röst och kommandon för att effektivt kunna tolka vad som sägs. Om systemet däremot är oberoende av talaren så behövs inte detta. Oberoende talsystem jämför talinmatningen (de ord som användaren säger) mot en samling redan förinspelade ord för att tolka vad användaren säger. Systemet försöker alltså hitta matchningar mellan de ord användaren säger och de förinspelade orden i systemets databas (Ayres & Nolan, 2006).

### **2.3. Röstbaserade system i fabriksmiljöer och andra kontexter**

I denna del av rapporten redogörs det för studier som genomförts om röststyrning inom fabriksmiljöer och i andra kontexter. Pires (2005) undersökte potentialen för röststyrda robotar inom tillverkningsindustrin. Liknande som Nass och Brave (2006) nämner, argumenterar Pires (2005) för att röststyrning är fördelaktigt i och med att tal är en naturlig form av kommunikation för oss människor, vilket gör det robust nog för att användas i mer kognitivt krävande situationer, till exempel inom industrier. Pires (2005) undersökte dessutom hur effektiv röststyrning är i både tysta respektive bullriga miljöer. När det kom till själva röststyrningen använde han sig av ett *automatiskt taligenkänningssystem* (eng. *Automatic Speech Recognition System*, eller *ASR*) i sina tester som var kopplat till ett par industriella robotar. I ett av experimentet skulle en robotarm, utifrån röstkommandon, röra sig från sin utgångsposition till ett objekt, greppa objektet, flytta det till en ny position och till sist återgå till sin utgångsposition. Detta test genomfördes i en tyst miljö, utan buller eller andra störningar som skulle kunna påverka systemets uppfattning av röstkommandona. I det andra experimentet skulle en robotarm, utrustad med en svets, svetsa en rak linje från en position till en annan. Det här exemplet genomfördes i en miljö med hög ljudnivå (ifrån svetsningen). Sammanfattningsvis förklarar Pires (2005) att detta röstbaserade system visade sig vara väldigt effektivt, även i en miljö med hög ljudnivå.

Fortsättningsvis, inom ämnet röststyrda robotar, utvecklade Kulyukin (2004) en assisterande robot som användaren kunde kommunicera med via tal. Kulyukin (2004) förklarar att med hjälp av en algoritm kunde fyra *typer* av interaktion uppnås (och valideras) mellan roboten och användaren i ett antal experiment. Fyra interaktionstyper som enligt Kulyukin (2004) är centrala för en *naturlig* röstkommunikation:

- *Kommandon*, roboten måste kunna förstå och utföra kommandon.
- *Disambiguering av mål*, för att roboten måste kunna tolka sina mål.
- *Introspektion*, roboten måste kunna svara på frågor om sina kunskaper och förmågor.
- *Instruktionsbaserad inlärning*, roboten måste kunna anpassa sig baserat på kommandon.

I en litteraturstudie av Koester (2001), med fokus på användarnas prestation vid användandet av röstbaserade system, tittade hon både på litteraturen kring en äldre typ av system: *Discrete Speech Recognition (DSR)* och en nyare typ: *Continuous Speech Recognition (CSR)*. Koester (2001) förklarar att DSR-system knappt används längre och att användare vid användningen av dessa system var tvungna att säga varje ord långsamt, tydligt och med pauser mellan orden för att systemet skulle kunna uppfatta dem korrekt. Med CSR-system däremot kan användare prata mer naturligt, och effektiviteten för dessa system har visat sig vara över 95% när det kommer till korrekt tolkning av talinmatning. Koester (2001) diskuterar även barriärer för framgång vid användningen av röstbaserade system. En av barriärerna är att användarna först måste lära sig de kommandon som används för att interagera med systemet, till exempel om användarna använder ett röstbaserat system för diktering skulle specifika kommandon för skiljetecken, stor bokstav, korrigering av fel, med mera, behöva memoreras av användare för att de effektivt ska kunna använda systemet.

Effektiviteten vid textredigering via röststyrning var något som Karat, Horn, Halverson och Karat (2000) bestämde sig för att undersöka närmare. De kom fram till att när det exempelvis gäller upptäckten och korrigeringen av fel vid diktering, såsom att systemet uppfattat fel ord, krävs det extra kognitiv ansträngning och tid för användaren att fixa det genom röststyrning i jämförelse med tiden det tar att åtgärda med mus och tangentbord. Det här beror troligtvis på att röststyrda gränssnitt inte är optimala när det kommer till textredigering, eller att användarna i undersökningen inte var medvetna om vilka tekniker eller kommandon som var möjliga/tillgängliga när det kom till att korrigera felen. I sin undersökning observerade de även att nya användare av denna typ av system oftare tenderade att använda sig av röststyrning för att rätta felen, vilket resulterade i sämre effektivitet (det tog längre tid att rätta felen). De mer erfarna användarna däremot sökte sig till multimodala sätt att korrigera texten, det vill säga tangentbord och mus, vilket troligtvis förklarar varför de var mer effektiva på det.

Ett annat problem associerat med röststyrning som Koester (2001) nämner har med rösttrötthet att göra. En del användare kanske inte märker av några problem överhuvudtaget relaterade till rösttrötthet, men risken för överansträngning av rösten existerar ändå. Kambeyanda, Singer och Cronk (1997) rapporterade exempelvis om fyra användare som utvecklat kroniska besvär efter långvarig (ett år) användning av ett DSR-system, vilket de förklarar troligtvis berodde på den monotona röst användningen som krävs vid diktering med denna typ av taligenkänningsystem. Det borde därmed inte vara något problem i dagsläget i och med att moderna taligenkänningsystem kan uppfatta naturligt tal, vilket Kambeyanda, Singer och Cronk (1997) menar inte borde medföra några direkta röstöveransträngningsproblem.

## ***2.4. User Experience Design och användarupplevelse***

Enligt ISO-standarden 9241-210 (2010) definieras *användarupplevelse* som: ”Personens perception och reaktion som ett resultat av användning av en produkt, ett system eller en tjänst” (SS-EN ISO 9241-2010, s. 3).

Användarupplevelse (eng. *User Experience*, eller *UX*) handlar om användarens upplevelse vid användningen av en produkt, ett system eller en tjänst (hädanefter används enbart termen *produkt*, men det syftas även på tjänster och system). Det berör till stor del hur människor känner och hur pass nöjda eller tillfredsställda de är vid interaktionen med en produkt, vilket kan påverkas av en rad faktorer. Användarnas intryck av produkten: hur den ser ut, känns, hur väl den gör det den är ämnad att göra och om den hjälper användarna att uppnå sina mål är några av de grundläggande faktorerna som påverkar användarnas upplevelse. När det kommer till den engelska termen för användarupplevelse: *User Experience*, läggs ibland ordet *Design* till (*UX-Design*, *UXD*) för att betona designaspekten när det kommer till att utforma användarupplevelsen (Preece, Rogers & Sharp, 2015/2016).

En annan term som ofta nämns och används inom fältet är *interaktionsdesign* (eng. *interaction design*) som myntades år 1984 av Bill Moggridge och Bill Verplank. Termen fick dock inte mycket uppmärksamhet förrän mitten på 1990-talet/tidigt 2000-talet i och med ett växande intresse bortom interaktiva produkters rena funktionalitet, till att estetiska designaspekter fick mer fokus (Lowgren, u.å.). Preece, Rogers och Sharp (2015/2016) förklarar att interaktionsdesign handlar om att utforma och skapa användarupplevelser som ”[...] förbättrar och utvidgar sättet som människor arbetar, kommunicerar och interagerar på” (Preece, Rogers & Sharp, 2015/2016, s. 29). De menar att en god användarupplevelse kan uppnås genom att maximera de positiva aspekterna av användarnas upplevelse vid interaktionen med en produkt, samtidigt som de negativa minimeras.

Preece, Rogers och Sharp (2015/2016) betonar också att det aldrig går att garantera en specifik användarupplevelse, det går enbart att designa för den. Det är användarna som i slutändan står för sin upplevelse, och designers kan enbart utforma produkter för att ge användare de bästa möjliga förutsättningarna för en specifik användarupplevelse. Det finns dock ett antal faktorer som går att ha i åtanke och utgå ifrån, som till stor del anses påverka användarupplevelsen, främst: ”[...] användbarhet, funktionalitet, estetik, innehåll, utseende och känsla, och den sensuella och emotionella attraktionen” (Preece, Rogers & Sharp, 2015/2016, s. 35).

I relation till emotionella aspekter, menar Hassenzahl och Tractinsky (2006) att affektion och känslor spelar en stor roll för användarupplevelsen. Däremot, precis som Preece, Rogers och Sharp (2015/2016) menar när de säger att det inte går att garantera en specifik användarupplevelse, förklarar Hassenzahl och Tractinsky (2006) att det är svårt att designa system som genererar specifika känslor eftersom det finns många bakomliggande faktorer som kan påverka dessa.

Hassenzahl och Tractinsky (2006) förklarar också att intresset för UX ökat de senaste åren som en opposition till det dominerande uppgifts- och arbetsrelaterade perspektivet på användbarhet. Hassenzahl och Tractinsky (2006) redogör vidare för tre olika perspektiv som de lyckats identifiera i litteraturen angående UX och potentiell framtida forskning inom fältet. I sin diskussion kring det första perspektivet förklarar de att tidigare forskning inom fältet *människa-datorinteraktion* främst fokuserat på *instrumentella mål*, såsom uppfyllandet av beteendemål och utförandet av uppgifter i en arbetsmiljö. Något som kommit att bli ifrågasatt och utmanat av de som argumenterar vikten av mer icke-instrumentella mål och hedoniska aspekter relaterade till användarupplevelsen (Hassenzahl & Tractinsky, 2006). Det är alltså viktigt att ha både de instrumentella och hedoniska aspekterna (som är av vikt för oss människor) i åtanke vid utformningen av produkter och tjänster för att kunna uppnå en högre standard när det kommer till användarupplevelsen. Hassenzahl och Tractinsky (2006) menar alltså att användarens upplevelse är resultatet av flertalet faktorer såsom produktens egenskaper och användarnas humör vid interaktionen.

Simpelt sagt påverkas användarupplevelsen av en rad komplexa faktorer som i sig påverkar varandra, vilket de menar (precis som nämnts tidigare) gör det svårt att designa en specifik upplevelse.

“UX is about technology that fulfils more than just instrumental needs in a way that acknowledges its use as a subjective, situated, complex and dynamic encounter. UX is a consequence of a user’s internal state (predispositions, expectations, needs, motivation, mood, etc.), the characteristics of the designed system (e.g. complexity, purpose, usability, functionality, etc.) and the context (or the environment) within which the interaction occurs (e.g. organisational/social setting, meaningfulness of the activity, voluntariness of use, etc.)” (Hassenzahl & Tractinsky, 2006, s. 95).

## **2.5. Användarcentrerad design, användarinvolvering och behov**

Det första steget i UXD-processen är vanligtvis explorativt, och det benämns ibland som den ”utforskande fasen”. Syftet med detta steg är att utforska de domän och den kontext som en produkt ska användas eller redan används i, samt *vem* som ska använda produkten. Det vill säga användarna, och vad de har för behov samt mål. I denna fas är det vanligt att utgå ifrån ett *användarcentrerat* tillvägagångsätt.

*Användarcentrerad design* går ut på att utveckla användbara system, produkter och tjänster genom att fokusera på användarnas behov genom hela utvecklingsprocessen. Genom implementeringen av användarcentrerade designmetoder i processen går det att uppnå bättre effektivitet vid användningen av produkter, skapa produkter som bättre tillgodoser användarnas behov samt en bättre användarupplevelse i allmänhet (SS-EN ISO 9241-210:2010).

Det har historiskt sett funnits en brist på kommunikation mellan utvecklare och användare. Det är inte helt ovanligt att utvecklare utgått ifrån sina egna idéer och åsikter vid utformningen av produkter, vilket kan resultera i ökad kognitiv stress för de faktiska användarna vid interaktionen, i och med att produkten inte är anpassad efter användarnas faktiska behov. Den här problematiken kan dock motverkas genom implementeringen av användarcentrerade metoder, vilket gör det möjligt för utvecklarna att enklare ta del av användarnas behov och därmed kunna skapa produkter som är bättre anpassade för att uppnå dessa behov (Obrist, Reitberger, Wurhofer, Förster & Tscheligi, 2011).

När det kommer till att utvärdera användarnas behov vid utvecklingen av en produkt är ett vanligt tillvägagångssätt att *involvera* användare tidigt i designprocessen, så kallad *användarinvolvering*. Kujala (2006) förklarar att vid involvering av användare i den användarcentrerade designprocessen kan användare anta flertalet olika *roller*; det kan röra sig om att användarna antar allt från en informativ till en konsulterande eller participativ roll. Kujala (2006) menar att genom involvering av användare i de tidiga faserna av designprocessen går det att få en bättre förståelse för användarnas arbetsprocesser och deras behov, vilket i slutändan kan resultera i en bättre slutprodukt som användarna är mer benägna att använda och nöjda med, vilket är en viktig bidragande faktor för en god användarupplevelse (Preece, Rogers & Sharp, 2015/2016).

Kontexten som produkten kommer användas i är också viktig att undersöka och förstå. Kujala (2006) menar att behov och kontext är sammanvävd. Kontexten begränsar nämligen möjligheterna för produkten och därmed bör även den utforskas och beaktas. Enligt ISO-standard 9241-210 består kontexten (eng. *context of use*) bland annat av användarna, uppgifterna de utför och den miljö de befinner sig i, vilket som nämnts tidigare är en viktig källa till information när det kommer till att specificera behov. Vidare, föreslår Kujala (2006) att behov kan utvinnas genom att identifiera problem som användarna upplever, eller som hindrar dem från att uppnå sina mål, och därefter utvärdera behov baserat på dessa problem.

“In case studies, it appeared clear that understanding the problems of the users was key to developing new products for them and that, when the problems were analysed, it was easy to develop new products to correct those problems and support users in their tasks. Thus we argue that user needs manifest themselves as either problems that hinder users in achieving their goals, or as opportunities to improve the likelihood of users achieving their goals.” (Kujala, 2006, s. 464).

## **2.6. Användarupplevelsen i fabriksmiljöer**

Enligt Wurhofer, Fuchsberger, Meneweger, Moser och Tscheligi (2015) lider det en brist på forskning kring människa-datorinteraktion i fabriksmiljöer och användarnas (arbetarnas) användarupplevelse i allmänhet. Wurhofer et al. (2015) förklarar att det kan bero på att undersökningar inom fabriksmiljöer medför en rad svårigheter och utmaningar, i och med att forskningen inte får komma i vägen för arbetsrutinerna. Obrist et al. (2011) nämner även detta och förklarar att en utmaning har med att arbetare inom fabriksmiljöer inte alltid har tid att delta i undersökningar på grund utav strikta tidsscheman.



Wurhofer et al. (2015) nämner en annan förklaring till bristen på litteratur inom området; de förklarar att forskningen som utförs och resultatet kan leda till konkurrensfördelar för företaget som forskningen utförs i samarbete med och att resultaten därmed kanske inte publiceras.

Reul (2009) föreslår att bristen på fokus kring användarupplevelsen inom fabriksmiljöer alternativt kan bero på att utvecklingen av industriella system ofta följer en mer traditionell utvecklingsprocess. Tillskillnad från konsumentprodukter som i allt större utsträckning följer en iterativ process och som vanligtvis lägger mer fokus på användarupplevelsen och involvering av användare i processen. En positiv användarupplevelse för arbetare i fabriker kan dock medföra flertalet positiva effekter (utöver att arbetarna är nöjda) exempelvis förbättrad arbetseffektivitet, reducerad inlärningstid samt en konkurrensfördel i allmänhet (Wurhofer et al., 2015). Däremot, till följd av början på den fjärde industriella revolutionen (benämnd ”Industri 4.0”), har det börjat läggas mer fokus på de individuella användarnas behov och upplevelse inom industrin (Jazdi, 2014).

Hur ser då användarupplevelse ut i nuläget inom fabriksmiljöer? I en undersökning av arbetarnas upplevelse i en fabrik (som tillverkade elektroniska mikrochips) noterade Obrist et al. (2011) att stress var den främsta negativa faktorn, vilket majoriteten av arbetarna nämnde var ett problem. Arbetarna nämnde även användbarhet och ergonomi som två viktiga faktorer i relation till deras användarupplevelse. Användbarhetsproblem var efter stress den faktor som påverkade deras upplevelse i störst utsträckning. Det visade sig även existera en koppling mellan användbarhetsproblem och stress; användbarhetsproblem visade sig vara den vanligaste orsaken till stress för arbetarna. Även sociala aspekter visade sig spela en viktig roll när det kom till användarupplevelsen (Obrist et al., 2011).

Wurhofer et al. (2015) genomförde en litteraturstudie där de gick igenom publicerad forskning som genomförts i den ovannämnda fabriken som tillverkade elektroniska mikrochips. De fokuserade på och undersökte information relaterande till användarnas upplevelse vid interaktionen med system i fabriksmiljön och identifierade tio faktorer (*UX-faktorer*) som observerats i olika studier inom fabrikskontexten hos användarna. *Perceived workload*, eller *uppfattad arbetsbelastning*, var faktorn som nämndes flest gånger i litteraturen. Denna faktor berör användarnas upplevelse av mängden kognitiv ansträngning behövd vid interaktionen med ett system för att lösa en uppgift. *Stress* var en annan vanligt förekommande faktor berörande användarnas uppfattade press och spänning relaterad till arbetet och interaktionen med system. Användarnas stressnivå påverkades bland annat av arbetsbelastningen samt en rad andra faktorer exempelvis: tidspress, skiftbyten och problem med system och utrustning. *Perceived usefulness*, eller *uppfattad användbarhet* är en faktor som berör huruvida användare upplever att ett system är användbart eller av nytta för dem, vilket påverkas av graden till vilket ett system förbättrar (eller försämrar) effektiviteten associerad med användarens utförande av arbetsuppgifter. *Trust*, eller *tillit*, berör användares upplevda tillit till att ett system kommer fungera som det är tänkt, vilket till exempel kan påverkas av antalet fel som uppstår, kvalitén på återkoppling från systemet, med mera (Se Wurhofer et al. (2015) för en mer detaljerad beskrivning av alla tio identifierade faktorer).

## 2.7. Problemprecisering

Syftet med detta arbete är att utforska behov relaterade till röststyrning inom fabriksmiljöer med målet att få en översiktlig bild av potentiella behov.

Utifrån ovannämnda mål för arbetet har följande frågeställning formulerats:

- *Vad finns det för generella behov relaterade till röstbaserade system inom fabriksmiljöer?*

För att avgränsa undersökningen gjordes valet att fokusera på två av de tre huvudsakliga användningsområdena för röstbaserade system som Ayres och Nolan (2006) redogör för: kommando och kontroll och diktering, samt ett tredje användningsområde som författaren av denna rapport själv identifierade: *återkoppling*, det vill säga röstbaserad återkoppling/information. Det sistnämnda användningsområdet (återkoppling) identifierades av författaren baserat på en beskrivning av användningsområden för röstbaserade system som Preece, Rogers & Sharp (2015/2016) gör. De nämner till exempel hur Apples *Siri* kan ge användarna röstbaserad information eller svara på frågor. Röståterkoppling är dessutom vanligt förekommande i flera moderna teknologier såsom GPS-system i bilar eller automatiserade telefonsystem och ansågs därmed vara ett intressant användningsområde att undersöka.

När det kommer till att utforma produkter och system inom UXD-fältet är det vanligt att utgå ifrån ett användarcentrerat tillvägagångsätt, och involvera användarna i processen där de kan anta en informativ roll (Kujala, 2006). Enligt ISO-standard 9241-210 (2010) är en vanlig aktivitet inom användarcentrerad design att undersöka användarnas behov associerade med den produkt, system eller tjänst som utvecklas. Enligt Kujala (2006) manifesterar sig behov ofta som problem som hindrar användare från att uppnå sina mål, eller som möjligheter att förbättra hur de kan uppnå sina mål inom en specifik kontext (användarnas mål i detta sammanhang definieras som vad det är de försöker uppnå vid utförandet av sina arbetsuppgifter eller interaktionen med system). Genom att försöka utröna användarnas behov (i detta arbete relaterade till röstbaserade system), och ha dem i åtanke under utvecklingsprocessen, går det som Kujala (2006) nämner att skapa produkter som i större utsträckning tillgodoser användarnas behov. System som användarna blir mer nöjda med samt mer benägna att faktiskt använda, vilket Preece, Rogers och Sharp (2015/2016) menar bidrar till att skapa en bättre användarupplevelse i överlag.

Att specifikt undersöka behov för röststyrning inom fabriksmiljöer är intressant i och med att vi människor är väldigt effektiva på att uppfatta och tolka meningen i röster och tal. Vi har genom evolutionens gång utvecklats till att bli experter på att utröna och tolka vad andra säger och menar (Nass & Brave, 2005). Dessutom, som Pires (2005) nämner, är tal en naturlig kommunikationsform för oss människor, vilket gör det till en fördelaktig form av kommunikation eftersom det kräver mindre kognitiv ansträngning och därmed låter användaren lägga mer fokus på utförandet av arbetsmomentet framför sig. Det finns därmed potential för röststyrning att användas som ett komplement inom fabriksmiljöer, och ifall det i framtiden skulle implementeras skulle det kunna förbättra systemens användbarhet, vilket också är en viktig faktor när det kommer till användarupplevelsen (Preece, Rogers & Sharp, 2015/2016).

Kontexten som en produkt ska användas i är också viktig att undersöka och förstå. Enligt ISO-standarden 9241-210 (2010) består kontexten (eng. *context of use*) av användarna, uppgifterna de utför och den miljö de befinner sig i. Kujala (2006) menar att behov och kontext är sammanvävd och att kontexten kan begränsa de praktiska möjligheterna, och till följd av detta kan även användarnas upplevelse begränsas. Därmed bör även kontexten utforskas och beaktas vid behovsanalyser. I relation till kontexten är det också intressant att undersöka hur användarnas upplevelse ser ut i nuläget, eftersom det utifrån användarnas nuvarande upplevelse är möjligt att utvärdera behov och möjligheter baserat på eventuella problem, frustrationer eller svårigheter som de upplever.

Sammanfattningsvis, är alltså målet med detta arbete att undersöka generella behov relaterade till röststyrning inom fabriksmiljöer. Detta kommer undersökas genom ett användarcentrerat tillvägagångssätt där användarna (arbetarna i den fabrik som ska undersökas) får anta en informativ roll. Resultatet, det vill säga insikten kring potentiella behov, skulle kunna utgöra grunden och informera fortsatt undersökning inom detta ämnesområde. Resultatet från fortsatt undersökning och en mer djupgående behovsanalys skulle därefter kunna användas vid en eventuell framtida utveckling av röstbaserade system för fabriksmiljöer. Genom att använda resultatet från behovsanalysen skulle system kunna utvecklas som i större utsträckning tillgodoser användarnas behov, vilket i slutändan borde bidra till en bättre användarupplevelse vid interaktion med systemen.

### 3. Metodologiska val

Enligt Berndtsson, Hansson, Olsson och Lundell (2008) är det viktigt att först klargöra vad målet med det projekt som ska genomföras är. Syftet med detta är dels för att klargöra och specificera vad projektet ämnar undersöka, dels för att avgöra vilka metoder som är bäst lämpade för projektet. Målet med detta arbete har huvudsakligen varit att undersöka behov i formen av möjligheter och begränsningar relaterade till röstbaserade system inom fabriksmiljöer. Efter det övergripande syftet med projektet är specificerat förklarar Berndtsson et al. (2008) att nästa steg går ut på att formulera ett antal delmål för projektet som ämnar bidra till att det övergripande målet uppnås. För varje delmål väljs det sedan en lämplig metod för att undersöka och uppnå målet. De huvudsakliga delmålen för detta projekts empiriska arbete var datainsamling och analys av insamlad data. Enligt Preece, Rogers och Sharp (2015/2016) är det en bra idé att kombinera flera datainsamlingsmetoder. Kombinationen av flera metoder menar de leder till att resultatet blir mer pålitligt, eftersom det möjliggör triangulering. För datainsamlingen valdes därför två kvalitativa metoder: *direktobservationer* och *semi-strukturerade intervjuer*. Insamlad data analyserades därefter både *deduktivt* och *induktivt* med hjälp av *affinitetsdiagram* med syftet att identifiera behov för röstbaserade system (se kapitel 4 för en genomgång av detta arbetets genomförande). Resultaten sammanställdes till sist och presenteras i denna rapportens resultatdel (se kapitel 5). Härnäst följer en redogörelse av de valda metoderna samt motiveringar för varför de valdes.

#### 3.1. Val av datainsamlingsmetoder

Det finns två huvudsakliga sorters data: *kvantitativ* och *kvalitativ*. Kvantitativ data kan enklast förklaras som den typ av data som det går att räkna statistiskt på, det vill säga data som består av numeriska värden, eller som går att omvandla till numeriska värden. Kvalitativ data däremot består av ord och meningar, beskrivningar, påståenden, respondenters känslor, karaktärsdrag, skildringar, med mera. Det är denna typ av data som ofta samlas in från exempelvis intervjuer. Kvalitativ data anses ibland vara mer nyanserad, men även subjektiv eftersom den oftast inte kan tolkas med statistiska mått, utan det är datainsamlaren själv som måste tolka vad den innebär (Patton, 2015). Det första steget inför en datainsamling är att specificera målet med datainsamlingen, det vill säga vilken data som ämnas samlas in, vilket kommer vara till hjälp vid avgörandet av vilka datainsamlingsmetoder som ska användas (Preece, Rogers & Sharp, 2015/2016). Detta arbete har till stor del varit utforskande och målet har varit att undersöka och utröna potentiella behov, varav data om följande aspekter valdes att samlas in:

- Möjligheter och begränsningar för röstbaserade system inom den aktuella fabriksmiljön.
- Den nuvarande användarupplevelsen i fabriken, däribland arbetarnas upplevelse associerade med datorsystemen de använder i sitt arbete, samt hur de upplever arbetsbelastningen och stressnivån.
- Den aktuella kontexten (fabriksmiljön, arbetarna och deras arbetsuppgifter).
- Deras tidigare erfarenheter och attityd till röststyrning.

Metoderna som valts och använts för datainsamling i detta arbete redogörs för härnäst.

### 3.1.1. Direktobservation

Observationer kan ske antingen *direkt*, genom att datainsamlaren är med och observerar kontexten och aktiviteter i realtid, eller *indirekt*, genom att aktiviteter spelas in och analyseras efteråt (Preece, Rogers & Sharp, 2015/2016). Observationer kan vidare utföras både i en *naturlig* miljö, det vill säga i den miljön som aktiviteterna av intresse vanligen brukar genomföras, eller i en *kontrollerad* miljö, det vill säga i ett labb. Både Kujala (2006) och Maquire och Bevan (2002) rekommenderar observationer när det kommer till att utforska kontexten vid behovsanalyser, vilket är den huvudsakliga motiveringen till varför denna metod valdes. Författaren kunde även själv se fördelar med att genomföra observationer; främst att de skulle kunna bidra med en överblick av kontexten och insikter som skulle kunna guida nästa del av datainsamlingen (intervjuerna)

En begränsning för observationerna inom fabriksmiljön identifierades dock. Begränsningen relaterade till det aktiva arbetet som utförs inom fabriken; observationerna fick inte störa de som arbetade och arbetsprocesserna. Därmed var det inte en möjlighet att i detalj observera de olika arbetsuppgifterna och samtidigt ställa frågor kring uppgifterna eftersom det skulle kunna störa arbetet. Observationerna skulle därmed bara kunna ske diskret och på avstånd, varpå beslutet gjordes att observationerna främst skulle genomföras för att undersöka själva miljön, med syftet att utforska om det fanns några direkt observerbara möjligheter och begränsningar för röstbaserade system inom fabriksmiljön.

### 3.1.2. Semi-strukturerad intervju

När det kommer till att arbeta användarcentrerat möjliggör intervjuer respondenter att anta en informativ roll, vilket (som nämndes i bakgrundsdel) är ett vanligt tillvägagångssätt när det kommer till att utröna behov. Intervjuer valdes även som en datainsamlingsmetod eftersom syftet med datainsamlingen till stor del var att samla in information om användarnas upplevelser, uppgifter och åsikter, vilket intervjuer är lämpliga för.

Intervjuer går ut på att samla in data från respondenter genom att huvudsakligen ställa frågor som antingen kan vara *öppna* eller *slutna*. Öppna frågor karaktäriseras av att de inte har några förutbestämda svarsalternativ. Istället ämnar de att uppmuntra respondenter till att fritt formulera sig, vilket vanligtvis genererar en större mängd kvalitativ data. Öppna frågor ger även respondenten större frihet att styra intervjun, tillskillnad från slutna frågor som vanligtvis inte ger respondenter någon möjlighet att utveckla sina svar. Däremot resulterar slutna frågor vanligtvis i mer strukturerad kvantitativ data som det är möjligt att utföra statistiska tester på, vilket är svårare med kvalitativ data (Berndtsson et al., 2008; Cooper, Reinmann, Cronin & Noessel, 2014).

Enligt Cooper et al. (2014) är utgångspunkten för etnografiska intervjuer antagandet att respondenten vet mer om det aktuella domän som frågorna ställs om än personen som ställer frågorna. Det är därmed fördelaktigt att ge respondenterna mer kontroll över intervjuerna. Cooper et al. (2014) menar även att risken finns att viktiga insikter missas om intervjun baseras på slutna frågor eftersom respondenterna kanske inte får möjligheten att dela med sig av dessa insikter. Samtidigt menar Cooper et al. (2014) att det är bra att ha några formulerade frågor att utgå ifrån baserade på de ämnen som intervjun ämnas täcka.

Preece, Rogers och Sharp (2015/2016) redogör fortsättningsvis för *semi-strukturerade intervjuer*; vid användningen av denna intervjustil utgås det från ett fåtal öppna frågor som respondenten får svara på och baserat på vad respondenten svarar kan lämpliga följdfrågor ställas. Denna intervjustil ger både respondenten och personen som ställer frågorna mer frihet och möjlighet att utforska intressanta ämnen i mer detalj. Baserat på detta valdes en semi-strukturerad intervjustil för intervjuerna som genomfördes i detta arbete. Det ansågs att denna intervjustil skulle göra det möjligt att utforska potentiellt intressanta områden, samtidigt som det skulle ge respondenterna viss frihet att själva styra intervjuerna.

Vidare, kan intervjuer precis som observationer, ske både på plats i respondentens naturliga miljö, det vill säga platsen där respondenten exempelvis utför aktivitet av intresse, eller i en neutral miljö, till exempel ett kontor eller ett labb. Enligt Cooper et al. (2014) är det fördelaktigt att genomföra intervjuer i respondenternas naturliga miljö eftersom respondenter kan få svårt att minnas hur de exempelvis utför uppgifter eller vilka objekt/verktyg de använder om intervjun sker på en plats avskild från var de utför uppgifterna. I detta arbete valdes det därför att intervjuerna skulle genomföras på plats i fabriken. Mer specifikt i fabriken entré som är i direkt anslutning till själva bearbetningsområdet i fabriken. Intervjuerna kunde dock inte genomföras vid respektive respondents arbetsstation, dels för att vissa arbetare jobbar vid många olika stationer/områden av fabriken, och dels för att detta skulle kunna störa det pågående arbetet.

Nästa steg i förberedelse av intervjuer går vanligtvis ut på att specificera den primära användargruppen, det vill säga den grupp av individer som det ämnas samla in data om, med syftet att rekrytera rätt deltagare för intervjuerna (Unger & Chandler, 2012). Patton (2015) förklarar att respondenter bör väljas baserat på undersökningens syfte. Det logiska valet var därmed att respondenterna skulle arbeta i den aktuella fabriken eftersom det var dessa individer och deras arbete, upplevelse och åsikter som var av intresse. Vid valet av möjliga respondenter att intervjua utgicks det därmed ifrån ett *kriteriebaserat urval*, dock var det enda kriteriet att respondenterna skulle arbeta i den specifika fabriken. Alternativt skulle urvalet kunna ses som ett *bekvämlighetsurval*, i och med antagandet att de flesta som rörde sig inom fabriken faktiskt arbetade där, och urvalet gjordes ”slumpmässigt” genom att passerande individer i fabriken frågades om de ville delta i en intervju (se fortsättningsvis kapitel 4).

När det kommer till att dokumentera respondenters svar går det exempelvis att föra anteckningar för hand på papper med penna. Däremot kan det bli svårt för personen som intervjuar att fokusera både på att anteckna och faktiskt lyssna på vad respondenten säger samt ställa meningsfulla följdfrågor. Ljudinspelning av intervjuer kan därmed vara fördelaktigt eftersom frågeställaren kan lägga mer fokus på respondenten och frågorna som ställs, och vid analysen kan exakt det som respondenterna svarade analyseras (Preece, Rogers & Sharp, 2015/2016). Valet gjordes därmed att intervjuerna skulle spelas in för att underlätta både vid intervjutillfället och vid den efterföljande analysen av intervjuvaren. Det finns dock ett antal etiska avvägningar relaterade till ljudinspelning som diskuteras vidare i kapitel 4.3.

### 3.2. Val av analysmetoder

I följande sektioner redogörs det för de metoder och tillvägagångsätt som valts för analysen av insamlad data i detta arbete.

#### 3.2.1. Deduktiv och induktiv analys

För analys av insamlad data valdes två metoder: *deduktiv analys* (eng. *deductive analysis*) och *induktiv analys* (eng. *inductive analysis*). De här två tillvägagångsätten valdes eftersom båda är vanligt förekommande och anses vara effektiva för analys av kvalitativ data (Patton, 2015). Författaren har även använt dessa analysmetoder tidigare och ansåg att de skulle lämpa sig vid analysen av både observationsdata och intervjudata eftersom denna data främst var kvalitativ.

Induktiv respektive deduktiv analys skiljer sig främst vid hur data kategoriseras. Vid deduktiv analys bearbetas data och kategoriseras i ett antal redan kända kategorier (Patton, 2015). Vid induktiv analys utgår det inte ifrån några specifika kategorier, istället bearbetas och analyseras data ofta i flera omgångar. Efter ett tag brukar mönster och teman gå att identifiera, varpå en kategorisering sker baserat på dessa mönster och teman som uppkommit (Patton, 2015). Ett mönster är en typ av deskriptivt fynd som upprepas flertalet gånger i materialet, medan ett tema kan ses som ett ämne eller kategori som återfinns. Patton (2015) beskriver och belyser skillnaden mellan teman och mönster med följande exemplifiering:

”The content analysis revealed a pattern of participants reporting being afraid when rappelling down cliffs and running river rapids; many also initially experienced the group process of sharing personal feelings as evoking some fear. Those patterns make dealing with fear a major theme of the wilderness education program experience.” (Patton, 2015, s. 541).

I exemplifieringen ovan tar Patton (2015) upp resultat från en undersökning av upplevelsen hos individer som deltagit i en överlevnadsutbildning. En återkommande aspekt Patton (2015) identifierat är att deltagarna i den undersökningen sagt sig uppleva rädsla vid diverse tillfällen, det vill säga ett mönster. Medan det huvudsakliga temat för denna överlevnadsutbildning kan sägas handla om deltagarnas hantering av rädsla.

#### 3.2.2. Affinitetsdiagram

Barnum (2011) redogör för en vanligt förekommande metod som används för att analysera och organisera kvalitativ data: *affinitetsdiagram* (eng. *affinity matching*). Metoden går ut på att insamlad data (som kan vara i formen av inspelningar, anteckningar, transkriberingar, med mera) först bearbetas med syftet att identifiera intressanta ”fynd”, exempelvis liknande svar, insikter, intressanta citat, med mera. Dessa fynd antecknas på exempelvis post-it-lappar som sätts upp på exempelvis en vägg eller whiteboardtavla. Lapparna läses sedan igenom, flyttas om och grupperas, ofta flertalet gånger. Efter ett tag brukar det gå att identifiera teman och mönster bland grupperingarna och det går då att börja klassificera grupperna baserat på de teman och mönster som identifierats. När alla som deltar i analysen är nöjda med placeringen av lapparna kan grupperingarna sedan namnges baserat på de mönster/teman som identifierats.

## 4. Genomförande

Under följande sektioner av rapporten redogörs det för hur det empiriska arbetet genomförts. Från planering och insamling till bearbetning och analys av data.

### 4.1. Direktobservationer

Inför observationer förklarar Preece, Rogers och Sharp (2015/2016) att det är viktigt att specificera syftet med observationen och vilken information som ämnas samlas in innan den genomförs. Syftet med observationerna för detta arbete var att samla in information som kan användas för att beskriva kontexten i fabriken, främst miljön, och för att undersöka observerbara eventuella möjligheter samt begränsningar för röststyrning. Observationerna syftade även till att ge en inblick i miljön som stöd och inspiration till utformningen av intervjufrågorna.

Preece, Rogers och Sharp (2015/2016) presenterar ett ramverk för direktobservationer som specificerar vad som ska observeras med syftet att styra fokuset för observationer. Ramverket de presenterar menar de fungerar bra som en utgångspunkt, men att det kan behöva modifieras beroende på syftet med studien. Följande ramverk utformades för observationerna och är delvis baserat på det ramverk som Preece, Rogers & Sharp (2015/2016) presenterar:

- *Rum:* Hur ser det fysiska rummet ut och hur är det utformat? Vad händer inom rummet?
- *Aktörer:* Vad går det att observera att aktörerna gör?
- *System:* Finns det några datorsystem och vad används de till?
- *Möjligheter:* Går det att observera några möjligheter för röststyrning?
- *Begränsningar:* Går det att observera några begränsningar för röststyrning?

Observationer genomfördes vid två tillfällen och pågick i cirka en timme vid varje tillfälle. Vid observationerna gick författaren igenom fabriken, observerade kontinuerligt aspekter utifrån samtliga frågor specificerade i det ovannämnda ramverket och förde anteckningar för hand i ett anteckningsblock. När det kom till observationer av datorsystem som användes, och möjligheter samt begränsningar för röststyrning i fabriken, fokuserade författaren av denna rapport på de tre tidigare nämnda användningsområdena, nämligen: Kommando och kontroll, diktering, samt återkoppling.



## 4.2. Semi-strukturerade intervjuer

Vid planeringen av intervjuerna specificerades det först vilken data de ämnade att samla in:

- Grundläggande demografisk data om respondenterna (ålder och hur länge de arbetat i den aktuella fabriken).
- Vad de arbetar med i fabriken och deras användarupplevelse relaterad till arbetsbelastning och stress.
- Om de använder sig av datorsystem i sitt arbete och hur de upplever användningen av systemen.
- Om de har någon tidigare erfarenhet av röstbaserade system, deras attityd till att använda röstbaserade system och om de kan se några möjligheter eller begränsningar med att använda det som ett komplement i sitt arbete.

### 4.2.1. Utformning av intervjufrågor och manus

Baserat på de ovan specificerade målen med intervjuerna utformades sedan intervjufrågorna (se *Bilaga 1* för det kompletta formuläret med intervjufrågor). Intervjuerna utgick ifrån en semi-strukturerad intervjustil med öppna frågor ämnade att samla in främst kvalitativ data. Eftersom intervjuerna skulle genomföras på respondenternas arbetstid förklarade handledaren på Volvo att intervjuerna inte fick ta för lång tid. Om intervjuerna blev för långrandiga fanns risken att respondenterna inte skulle ha tid att besvara frågorna, och författaren och handledaren på Volvo kom därmed överens om att fem till tio minuter skulle vara en lämplig tid för intervjuerna. Tiden att svara på varje fråga estimerades först av författaren och testades sedan i pilottestet (se sektion 4.2.2.) som bekräftade att mängden frågor och tiden det tog att besvara dem var lämplig.

För enkelhetens skull (inför genomförandet) delades intervjufrågorna in i fyra ”delar” baserat på vilken data de ämnade samla in. Del 1 ämnade samla in demografisk data om respondenterna, men även för att ”värma upp” själva intervjun. Enligt Preece, Rogers & Sharp (2015/2016) är syftet med att samla in grundläggande demografisk data om respondenterna att kunna sätta svaren i ett sammanhang och möjligtvis kunna utröna skillnader i svaren. Om respondenter har olika åsikter kanske detta beror på exempelvis deras ålder eller hur länge de arbetat inom fabriken, vilket därmed går att analysera baserat på de demografiska uppgifterna. Del 2 ämnade samla in data om respondenternas arbetsuppgifter och deras användarupplevelse relaterad till sitt arbete. Del 3 berör användningen av datorsystem och respondenternas upplevelse associerad med användningen. De sista frågorna, del 4, berör respondenternas tidigare erfarenhet, upplevelse och attityd till användningen av röstbaserade system, samt om respondenterna kunde se några möjligheter eller begränsningar för användning av röstbaserade system vid deras arbete inom fabriken. Respondenterna frågades om deras tidigare erfarenheter av röststyrning samt deras attityd till att använda det i sitt arbete eftersom det är aspekter som både kan medföra begränsningar och påverka användarupplevelsen ifall att röstbaserat system skulle utvecklas. Om användarna till exempel är negativt inställda kan det medföra illvilja mot att använda ett eventuellt framtida system, vilken troligtvis resulterar i en negativ användarupplevelse.

Undersökning av attityden kan även medföra insikter kring möjligheter och begränsningar; om användare till exempel har en negativ attityd kan det bero på att de kan se vissa begränsningar som man själv inte är medveten om eller kan observera.

Följande intervjufrågor (fråga 5, 6, 6.1, 8 och 9) ämnade samla in information om respondenternas användarupplevelse:

5. Hur upplever du arbetsbelastningen?
6. Hur upplever du stressnivån i ditt arbete?
  - 6.1 **Om de upplever någon form av stress:** Kan du beskriva vad du gör när du upplever denna stress?
8. Hur upplever du det är att använda systemet/en?
9. Har du upplevt några problem med systemet/en?

Frågorna baserades på och ämnade undersöka de fyra vanligast förekommande UX-faktorer som Wurhofer et al. (2015) identifierat inom fabriksmiljöer: *uppfattad arbetsbelastning*, *stress* samt *uppfattad användbarhet* och *tillit* till system. Fråga 8 och 9 berör inte uppfattad användbarhet och tillit specifikt men de ämnar fånga aspekter som relaterar till och påverkar användbarhet och tillit, exempelvis upplevda problem med systemet.

Vid observationerna noterades det bland annat att användarna i fabriken tog del av information från datorsystem samt matade in diverse information i systemen. Följande intervjufrågor (fråga 7.2 och 7.3) utformades därmed för att undersöka detta vidare:

- 7.2. Tar du emot någon form av information via systemet? **Om ja:** Kan du beskriva detta närmare?
- 7.3. Förs det anteckningar eller matas in information av någon form i systemet?

Till sist utformades även ett manus som skulle läsas upp för respondenterna innan intervjun påbörjades och mellan de olika delarna av intervjun (se *Bilaga 1*). Manuset ämnade att konsistent förmedla samma information till varje respondent och för att säkerställa att ingen viktig information missades att förmedlas. Manuset innehöll främst information om studiens syfte, respondenternas rättigheter, aspekter relaterade till hanteringen av deras svar på frågorna och deras anonymitet.

#### 4.2.2. Pilottest

En pilottest är ett test av studien och de metoder som planeras att användas och syftar till att säkerställa att undersökningen och metoderna är praktiskt tillämpbara, samt identifiera eventuella övriga problem. Hela undersökningen genomförs som planerat fast med bara en eller ett fåtal deltagare och därefter reflekteras det över vilka eventuella problem med utformningen av studien eller metoderna som observerats under testet. Om några problem identifierats åtgärdas dem och eventuellt genomförs ytterligare ett pilottest innan den faktiska studien genomförs.

Efter planeringen av intervjun och utformningen av intervjufrågor var klar genomfördes ett pilottest. Testet genomfördes på plats i fabriken där en slumpmässig arbetare frågades om hen ville delta i en undersökning och svara på några frågor.

Längden på pilottestet var ungefär tio minuter, vilket bekräftade att intervjuerna troligtvis skulle hamna inom det bestämda tidsspannet på fem till tio minuter.

Efter själva intervjun var slutförd frågades respondenten om hen upplevde att några frågor var svåra att förstå eller besvara, samt om det var något annat som var otydligt relaterat till intervjun. Respondenten indikerade att hen inte upplevde några problem eller svårigheter med att förstå eller besvara frågorna. Författaren upplevde heller inte några problem med intervjun eller frågorna och därmed beslutades det att undersökningen var redo att genomföras.

#### **4.2.3. Genomförande av intervjuer**

Intervjuerna genomfördes även de på plats i fabriken. Författaren (samt en annan student) placerade sig i entrén till fabriken (som är i direktanslutning till själva fabriksområdet) där många arbetare passerar. Författaren hälsade på passerande arbetare och frågade om de ville delta i en intervju och besvara några frågor. Som incitament för att delta bjöds det även på bullar till de som valde att delta. Vid bemötandet frågades arbetarna om de arbetade i den specifika fabriken för att säkerställa att de uppfyllde urvalskriteriet. Intervjuerna genomfördes därefter och spelades in på författarens mobiltelefon (efter respondenterna gett sitt godkännande till detta). Totalt intervjuades 14 personer och intervjuerna var cirka mellan fem till tio minuter långa. Författaren utgick ifrån de frågor som formulerats, men beroende på respondenternas svar improviserades följdfrågor under intervjuernas gång, vilket gav mer frihet till respondenterna att utveckla sina svar, samt möjlighet för både respondenterna och författaren att styra intervjuerna i större utsträckning. Följdfrågorna blev därmed till stor del *sonderande*, det vill säga de undersökte vissa ämnen mer djupgående (Preece, Rogers & Sharp, 2015/2016).

När det kommer till kvalitativ datainsamling kan det vara svårt att avgöra när tillräcklig mängd data är insamlad. Datainsamlingar kan i princip fortgå under en oförbestämd tid. Det går aldrig med 100% säkerhet säga all relevant data är insamlad vid något tillfälle. Det går däremot att nå *informationsmättnad* vilket innebär att den data som samlas in efter ett tag inte upplevs bidra med mycket ny eller unik information i relation till den data som redan samlats in. Mot slutet av intervjutillfället i detta arbete började viss informationsmättnad att uppnås. Respondenternas svar uppfattades som likartade, och trots att respondenternas svar angående deras arbetsuppgifterna skiljde sig åt fattades beslutet att tillräcklig information hade samlats in för analys. Valet att avsluta intervjuerna gjordes även av tidsskäl. Cirka två timmar med ljudinspelningar hade samlats in, och av tidigare erfarenhet visste författaren att detta skulle ta lång tid att transkribera och analysera.

### ***4.3. Etiska aspekter relaterade till insamling och hantering av data***

Vid varje intervjutillfälle informerades respondenterna allra först om deras rättigheter, däribland information om deras anonymitet, vilka som skulle ta del av svaren, vad syftet med studien var, med mera. Informationen de fick var formulerad sedan innan och lästes upp för varje respondenterna innan själva intervjun påbörjades för att säkerställa att de alla fick ta del av samma information och för att ingen viktig information skulle missas (se *Bilaga 1* för den information varje respondent fick uppläst för sig). Om respondenternas anonymitet är viktig nämner Berndtsson et al. (2008) att hanteringen av insamlat material måste övervägas noga, till exempel ljudinspelningar. Vid diskussion med handledaren på högskolan kom författaren av denna rapport fram till att ljudinspelningarna från intervjuerna skulle vara den känsligaste informationen associerad med respondenternas anonymitet och därmed beslutades det att författaren skulle fråga handledaren på Volvo om de skulle ha något intresse av inspelningarna. Det visade sig att Volvo inte hade något intresse av inspelningarna i sig, utan de var mer intresserade av slutresultatet. Det bestämdes därmed att efter arbetet var slutfört skulle inspelningarna raderas och att respondenternas svar enbart skulle ges ett id-nummer för att säkerställa att deras svar inte skulle kunna kopplas till dem personligen. Berndtsson et al. (2008) diskuterar vidare kring huruvida det är nödvändigt att samla in demografisk data om respondenter. De menar att ifall den demografiska informationen går att koppla till en specifik respondents svar när de presenteras utgör det en risk för respondenternas anonymitet. Även om inte några namn nämns menar de att om tillräckligt många detaljer om respondenterna framgår finns risken att deras identiteter går att utröna:

“To what extent there is a need to characterise the interviewees, and what aspects of their background that are of significance to the study. This issue is closely related to the issue of anonymity in your study. In writing your report there are a number of potential pitfalls. For example, even though the name of a specific individual is not mentioned, it might be very easy for colleagues in the organisation to identify that person.” (Berndtsson et al. (2008), s. 71).

Den demografiska data om respondenterna som samlats in vid intervjuerna i detta arbete är respondenternas ålder och hur många år de arbetat i den specifika fabrikslokalen. I och med att det är relativt många som arbetar i den aktuella fabriken är det troligtvis svårt för någon att utröna specifika respondenter identitet baserat enbart på deras ålder och hur länge de arbetat i fabriken. Det är dock inte omöjligt, speciellt om någon är bekant med en respondent. Därmed valdes det att denna data inte skulle presenteras i rapporten, utan enbart användas av författaren vid behov för att exempelvis jämföra svar mellan olika åldersgrupper eller liknande.

Respondenternas svar angående deras arbetsuppgifter i fabriken ansågs också vara känslig relaterad till deras anonymitet av samma anledning som nämns ovan; ifall någon är bekant med respondenten, eller om respondenten utför arbetsuppgifter som enbart ett fåtal anställda gör, skulle det gå att identifiera vem respondenten är. Därför gjordes valet att denna information heller inte skulle presenteras i rapporten, åtminstone inte på ett sätt så det skulle kunna identifiera de exakta arbetsuppgifterna (i kapitel 5.2.1. av resultatdelen diskuteras detta vidare).

Vid vissa tillfällen kan det krävas att ett skriftligt kontrakt för informerat samtycke signeras innan genomförandet av intervjuer. Preece, Rogers och Sharp (2015/2016) menar dock att ett skriftligt kontrakt för samtycke inte alltid är nödvändigt när det kommer till datainsamling om det redan finns ett existerande kontrakt mellan personen som samlar in data och företaget som beställt datainsamlingen. Följande citat från Preece, Rogers och Sharp (2015/2016) tydliggör resonemanget på följande sätt:

”Ta t.ex. en situation där en konsult är inhyrd för att samla in data från ett företag för att kunna fastställa de behov som ställs för ett nytt interaktivt system som ska stödja tidsrapportering. De anställda i det här företaget skulle utgöra användarna av systemet. Därför förväntar sig konsulten att få tillgång till de anställda för att samla in data om tidsrapportsaktiviteten. Dessutom skulle företaget förvänta sig att de anställda samarbetar i den här aktiviteten. I det här fallet finns det redan ett kontrakt som täcker datainsamlingsaktiviteten och därför är det mindre troligt att det krävs ett formulär för informerat samtycke.” (Preece, Rogers & Sharp (2015/2016), s. 288-290).

Det ansågs därmed inte nödvändigt att ta fram ett skriftligt kontrakt för samtycke vid intervjuerna eftersom det inför detta arbete redan etablerats ett kontrakt mellan författaren av denna rapport och Volvo som säger att företaget *kan* välja att ge författaren av denna rapport nödvändig information för att genomföra arbetet, men att handledaren på företaget måste godkänna informationen innan den publiceras. Vid diskussion med handledaren på Volvo gavs det godkännande att genomföra intervjuer med anställda i fabriken, och det ansågs därmed räcka att respondenterna gav sitt muntliga godkännande för deltagande vid intervjutillfället.

#### ***4.4. Analys av observationsdata***

Insamlad data från observationerna bestod av ett antal sidor med anteckningar från observationstillfällena. Preece, Rogers och Sharp (2015/2016) förklarar att den inledande bearbetningen av denna typ av data går ut på att renskriva och eventuellt komplettera anteckningarna med ytterligare kommentarer eller insikter i efterhand, vilket gjordes direkt efter respektive observationstillfälle.

Det valdes att ett affinitetsdiagram skulle skapas och användas som stöd vid denna analys. Observationsanteckningarna lästes igenom flertalet gånger och ”fynd” identifierades (till exempel en rumslig beskrivning) som skrevs ner på post-it-lappar. Om ett ramverk används vid observationer förklarar Preece, Rogers och Sharp (2015/2016) att det sker en slags strukturering av data redan vid insamlingstillfället. Som nämndes tidigare användes det ett ramverk för observationerna som guidade datainsamlingen, och insamlad observationsdata antogs därmed kunna klassificeras baserat på de kategorier som ramverket bestod av: rum, aktörer, system samt möjligheter och begränsningar för röststyrning. Insamlad observationsdata analyserades därmed deduktivt (se kapitel 3.2.1. om induktiv respektive deduktiv analys), och vid skapandet av affinitetsdiagrammet placerades och klassificerades fynden (nedskrivna på post-it-lappar) baserat på de olika kategorierna i ramverket för observationerna. Resultatet av denna analys blev ett antal kluster-kategorier bestående av observationsdata som sedan kunde sammanställas och presenteras (se sektion 5.1. i resultatdelen).

#### ***4.5. Analys av intervjudata***

Det första steget i analysen gick ut på att transkribera ljudinspelningarna från intervjuerna för att underlätta vid analysen.

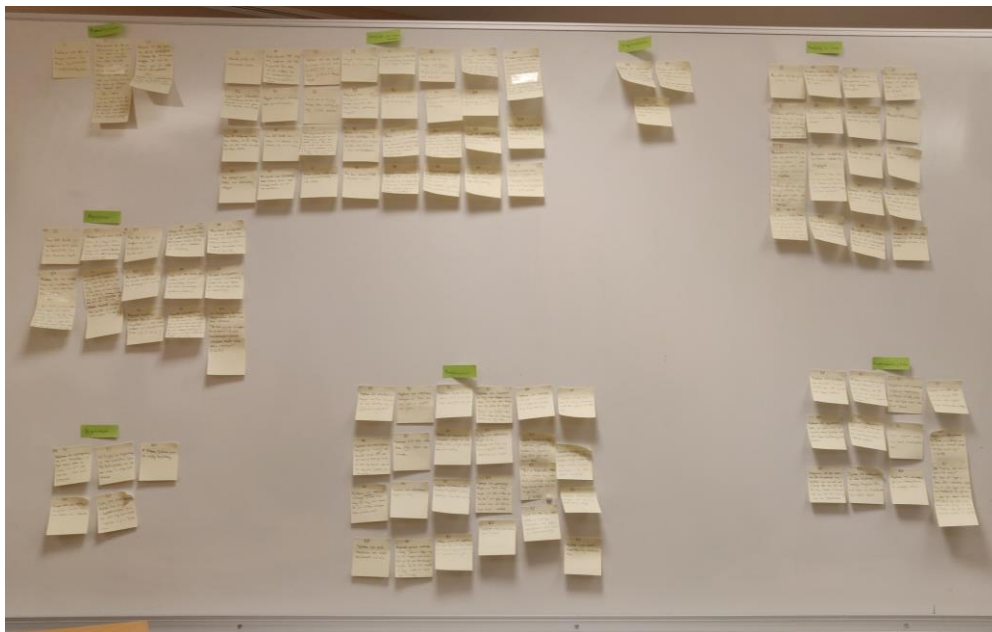
Patton (2015) förklarar att det första steget vid bearbetning och analys av insamlad data brukar gå ut på att försöka klassificera de primära identifierbara mönstren. Som nämndes i kapitel 4.2. ämnade intervjuerna att samla in följande data och intervjufrågorna formulerades därefter:

- Grundläggande demografisk data om respondenterna (ålder och hur länge de arbetat i den aktuella fabriken).
- Vad de arbetar med i fabriken och deras användarupplevelse relaterad till arbetsbelastning och stress.
- Om de använder sig av datorsystem i sitt arbete och hur de upplever användningen av systemen.
- Om de har någon tidigare erfarenhet av röstbaserade system, deras attityd till att använda röstbaserade system och om de kan se några möjligheter eller begränsningar med att använda det som ett komplement i sitt arbete.

Därmed antogs det vid början av analysen att insamlad intervjudata inledningsvis borde kunna klassificeras baserat på ovan nämnda mål för intervjuerna, det vill säga svaren skulle inledningsvis kunna analyseras och klassificeras deduktivt.

Det valdes att ett affinitetsdiagram skulle skapas även för analysen av denna data. Transkriberingarna lästes igenom och fynd antecknades på post-it-lappar. Fynden bestod huvudsakligen av respondenternas svar på frågor, citat och sammanfattningar av vad respondenterna svarat. Det sker därmed en viss tolkning av respondenternas svar redan i det här steget eftersom respondenternas ordagranna svar sammanfattades och tolkades till viss del när de antecknades på lapparna.

Först placerades lapparna beroende på vilket av de ovannämnda målen för datainsamlingen de berörde: grundläggande demografisk data, deras arbete i fabriken, användarupplevelse, användning av datorsystem, respondenternas attityder till röststyrning och vilka möjligheter samt begränsningar för röststyrning de kunde se i sitt arbete. Resultatet blev ett antal kluster med lappar kategoriserade utefter ovannämnda kategorier (se *Figur 1*). Därefter analyserades fynden ytterligare en gång, fast denna gång inbördes kategori-klustren, och induktivt istället för deduktivt. Vid denna analysomgång utgicks det inte från några förbestämda kategorier, utan lapparna analyserades och placerades om (inom respektive kategori-kluster) baserat på mönster och teman som kunde identifieras. Till sist, kunde dessa nya kluster klassificeras och namnges baserat på identifierade mönster och teman. Till exempel kunde klustren inom kategorin ”Attityder till röststyrning” klassificeras och namnges som exempelvis ”Positiva” och ”Negativa” attityder.



*Figur 1.* Ett av de affinitetsdiagram som skapades för att analysera insamlad data i detta arbete.

## 5. Resultat

I detta kapitel presenteras det analyserade och sammanställda resultatet från det empiriska arbetet som genomförts. Först redogörs resultaten från direktobservationerna följt utav resultaten från de semi-strukturerade intervjuerna. Se vidare kapitel 6 för en redogörelse och diskussion kring de behov som identifierats baserat på följande resultat.

### 5.1. Resultat från direktobservationerna

Observationerna ämnade främst att ge en uppfattning av kontexten (främst själva fabriksmiljön) och för att potentiellt bidra med information och insikter som kunde vara användbara vid utformandet av intervjufrågor. Kontexten kan som nämnts tidigare i rapporten även ge insikter om möjligheter och begränsningar, vilket var den andra huvudsakliga anledningen till att observationer valdes att genomföras.

#### 5.1.1. Rumsliga observationer

I entrén till fabriken stod det på en skylt vilka komponenter som bearbetas i fabriken, nämligen: *cylinderblock*, *cylinderhuvud*, *kamaxlar*, *kugghjul*, *svänghjul* och *vevaxlar*. Bearbetningen av komponenter såg till större delen ut att vara automatiserad och utfördes av maskiner/robotar.

I vissa utrymmen i fabriken förvaras komponenter, men den huvudsakliga ytan ser ut att täckas av "bearbetningsenheter". Dessa bearbetningsenheter består av maskiner/robotar som utför diverse bearbetningsmoment på komponenter. De flesta av dessa enheter är avskilda eller inhägnade bakom skydd.

Fabriken har god belysning, främst lysrörsbelysning, och det är högt i tak. I taket går det flera parallella "spår" som en maskin med en höj- och sänkbar robotarm åker fram och tillbaka på och det verkar som att den (åtminstone till viss del) förflyttar de olika komponenterna mellan bearbetningsenheter. Det finns markerade gångar och områden för var personer ska gå och var de behöver vara uppmärksamma på eventuella fordon (till exempel gaffeltruckar). Ljud/-bullernivån i fabriken upplevdes som varierande, det vill säga den upplevdes skilja sig mellan de olika delarna av fabriken.

#### 5.1.2. Observationer kring aktörer

Operatörer och andra arbetare använder hörselskydd och en del av dem har privata hörlurar/öronsnäckor (troligtvis för att lyssna på musik eller liknande). Operatörerna ser främst ut att utföra service, kalibrering och programmering av maskiner/bearbetningsenheter. De använder sig av både stationära-datorer och system inbyggda i maskinerna/bearbetningsenheterna. Andra arbetare ser ut att kontrollera eller justera komponenter vid diverse arbetsbänkar.

Det observerades även att arbetare lyfte och transporterade komponenter och andra objekt med hjälp av motordrivna gaffeltruckar. Dessa truckar rörde sig inom egna körbanor där de har företräde och som gående arbetare ej ska vistas inom.



Det finns dessutom kontor inom fabrikslokalen där arbetare ser ut att utföra andra typer av arbete framför datorer, möjligtvis administrativt arbete. Författaren av denna rapport kom i kontakt med en individ under en av observationerna som arbetade som "ledare" inom fabriken. Denna arbetsbefattning förklarade individen innefattade ett personalansvar som till stor del gick ut på att delegera arbetsuppgifter och se till att arbetare fanns på alla stationer som hen var ansvarig för samt se till att det som behöver bli gjort blir klart i tid.

### **5.1.3. Observationer kring användningen av system**

I anslutning till maskinerna/bearbetningsenheterna finns det diverse "datorskåp" med exempelvis en display och knappar som ger information om bearbetningen och tillåter operatörerna att programmera maskinen.

Det finns även "vanliga" stationära-datorer med mus och tangentbord. Vid en närmare observation av en del datorer ser det ut som de används för planering/dokumentering av något slag. Det ser även ut som arbetarna kan ta/tar emot information/instruktioner via datorerna om arbetsuppgifter de ska utföra.

Vid en datorstation fanns det en laserskanner och en lapp med QR-koder som, enligt en utskrivna beskrivning, användes för att navigera till webbsidor (genom scanning av QR-koderna). Webbsidorna det fanns QR-koder för var länkar till webbsidor med arbetsinformation, tidsinformation, produktionsdata, med mera.

Det finns även större skärmar placerade runt om i fabriken, vid olika arbetsstationer och bearbetningsenheter. Informationen på dessa skärmar varierar men det verkar som de visar diverse information om fabriken och bearbetningen, exempelvis el- och energiförbrukningen, aktuella nyheter, med mera.

### **5.1.4. Observerade potentiella möjligheter för röststyrning**

- Skulle eventuellt gå att ta emot arbetsinstruktioner via röst samtidigt som arbetet utförs.
- Om det förs anteckningar i systemet; skulle dessa möjligtvis gå att föra via diktering.
- Skulle kanske vara möjligt att programmera/kalibrera bearbetningsenheter via röstkommandon (om det inte är för avancerade steg involverade).

### **5.1.5. Observerade eventuella begränsningar för röststyrning**

- Ljudnivån i fabriken skulle eventuellt kunna medföra problem, både när det kommer till återkoppling via röst och talinmatning.
- Aktörerna använder hörselskydd; kan medföra hinder för att ta emot information eller ge kommandon.
- Röststyrningen skulle eventuellt kunna störas om flera aktörer jobbar i närheten av varandra och använder röstkommandon eller får återkoppling samtidigt.

Till följd av den observerade bullernivån bestämde sig författaren spontant för att genomföra ett antal enkla test med dikteringsfunktionen i sin smartphone vid ett av observationstillfällena. Syftet med detta test var att undersöka och få en uppfattning av bullernivåns eventuella påverkan på tolkningen av talinmatning. Författaren testade att tala in en mening vid tre olika platser i fabriken och det visade sig att talinmatningen tolkades relativt effektivt, trots bullernivån. Detta test var inte planerat och det går inte att dra några faktiska slutsatser baserat på det utan vidare tester skulle behöva genomföras först. Författaren anser dock att det är en intressant observation att nämna i och med att om en simpel mikrofon inbyggd i en smartphone kan tolka talinmatningen relativt bra så borde mer avancerade mikrofoner kunna göra det ännu bättre.

## **5.2. Resultat från de semi-strukturerade intervjuerna**

Totalt intervjuades 14 personer som arbetade inom fabriken, dock missades vissa frågor att ställas till vissa respondenterna vid några intervjuer, antingen av misstag eller för att det inte var passande beroende på hur intervjuerna ”flöt på”. Hur många respondenter som frågats respektive fråga tydliggörs inom parantes i kommande underrubriker.

Respondenterna benämns med bokstaven ”R” följt av en siffra för att kunna skilja mellan respondenterna och deras svar (till exempel är R2 = *respondent 2*).

### **5.2.1. Respondenternas arbetsuppgifter samt dator- och systemanvändning i fabriken**

Av sekretesskäl går det inte att specificera exakt vilka arbetsuppgifter och system eller programvaror respondenterna nämnt att de utför/använder eftersom, precis som Berndtsson et al. (2008) nämner, skulle det kunna gå att utröna vem respondenten är om denna typ av data från intervjuerna tas med i rapporten. Det som går att redogöra för dock är att respondenterna utförde en rad olika arbetsuppgifter och alla använde sig av ett eller flera datorsystem i sitt arbete. Respondenterna tar både del av information via de olika systemen, exempelvis arbetsinstruktioner, och matar in information i systemen, till exempel information om vad de utfört för service på en maskin, felbeskrivningar, med mera.

### **5.2.2. Respondenternas nuvarande upplevelse vid dator- och systemanvändning**

Sju respondenter (7 av 12 respondenter som frågades) indikerade att de i allmänhet upplever interaktionen med datorerna och systemen som de använder i sitt arbete som positiv (R1, R4, R5, R6, R8, R10 och R14).

Fyra respondenter (R1, R4, R6 och R8) svarade att de inte upplevt några direkta problem med systemen, eller att de inte kunde minnas några. R5 svarade följande på frågan om hur hen upplever användningen av det system hen använde mest i sitt arbete: ”Jag tycker det är bra, men det finns ju alltid smågrejer man irriterar sig på.”. R10 svarade att användningen av systemen i överlag är ”väldigt enkelt”.

Fem respondenter (5 av 12) hade upplevt viss problematik med systemen de använder (R2, R3, R7, R9 och R13). R2 upplever användningen av ett system som ”OK ” men nämnde även att systemet kunde vara lite ”segt” ibland. R2 tyckte även att EVO-processen är ”knölig” (som respondenten förklarade är en form av felrapporteringsprocess):

”[...] vi har något som heter EVO-process eller EVO [...] EVO det är nånting.. när nånting stannar ute i produktion så ringer man underhåll och det kommer dit nån, en elektriker eller mekaniker, som ska felsöka för att få igång maskinen. Och när man gör det här så använder man en.. det kallas EVO, man skriver alltså ner vad som är fel och vad man kan motverka nästa gång och ah liknande. Men den är lite knölig [...]” (Respondenten förklarar hur EVO-processen går till i nuläget). ” [...] men det tar ett tag att sätta sig in i den här processen, hur man gör. Vi hade velat haft att du klickar, och så skriver du i och sen tabbar vidare till nästa.” (R2)

Vid frågan om hur R3 upplevde användningen av EVO-systemet svarar hen: ”Det är inte så dåligt är det inte, men det är lite stelbent ibland liksom att hitta tycker jag.”

Författaren frågar då: ”Vad är det som kan vara svårt att hitta?”

R3 svarar: ”Asså allt ligger ju i samma.. Alla reservdelar ligger där, alla jobb ligger där, alla fel ligger där som vart historiskt så det blir ju som en enorm databas och ibland blir det lite svårt att skilja ur och filtrera vad det är man försöker leta upp.”

R7 säger att systemet hen använder i överlag fungerar bra men att det kan uppstå frustrationer ibland. R9 upplever användningen av TNT-systemet (som används för att spåra individuella komponenter) som ”helt okej”, men att det finns rum för förbättring av det. R13 frågades om hen upplevde några problem med systemen hen använder och svarade då att hen upplever det som problematiskt med nya systemuppdateringar när hen precis lärt sig systemet som uppdateras: ”Njaa, ibland när dom håller på och ändrar liksom uppdateringar och nya system när man precis hållt på att lära sig det gamla som är.”

### **5.2.3. Respondenternas nuvarande upplevda arbetsbelastning**

Tre respondenter (3 av 14 respondenter som frågades) upplever arbetsbelastningen som låg (R3, R4 och R13). R4 svarade att hen upplever arbetsbelastningen som: ”Väldigt bra”. Även R3 upplever att arbetsbelastningen är ”bra” och förklarade att tiden de har på sig att utföra uppgifter är tillräcklig: ”Den är ganska bra tycker jag, man har tid. Står det att det tar två timmar på ett jobb så tar det två timmar ungefär. Inte mycket mer, inte mycket mindre.”

Tre respondenter (3 av 14) respondenter upplever arbetsbelastningen som ”lagom” eller ”OK” (R1, R6 och R8). R6 förklarade att arbetsbelastningen var ”helt okej” och att det sällan var ”för lite” att göra, men upplevde inte något ”högt tryck” i arbetet.

Fem (5 av 14) respondenter upplever arbetsbelastningen som varierande, och att den skiljer sig åt från tid till tid (R5, R9, R10, R11 och R14). R11 förklarar även att den varierar beroende på vilken maskin (syftar på bearbetningsmaskin- eller station) hen arbetar på.

”Det beror på från tillfälle till tillfälle. Hur bråttom det är om det är en massa folk som står och väntar på en.” (R5).

”Den är väldigt vågig, det toppar och dalar [...] Det kan bli att det kommer smällar då det blir mycket en dag, mindre en annan dag.” (R9).

Tre respondenter (3 av 14) upplever arbetsbelastningen som hög (R2, R7 och R12). R7 förklarade att arbetsbelastningen var varierande men ofta hög och att det var mycket att göra.

#### **5.2.4. Respondenternas nuvarande upplevelse av stress**

Sex respondenter (6 av 14 respondenter som frågades) upplever stressnivån i sitt arbete som låg (R1, R2, R3 R4, R8 och R13). R2 förklarar att även fast det finns mycket att göra (hög arbetsbelastning) upplever hen inte mycket stress i sitt arbete. R3 upplever att stressnivån är låg men förklarar samtidigt att hen kan stressa upp sig själv om något börjar gå dåligt, men att det snarare beror på pressen hen sätter på sig själv, och menar att det inte beror på att arbetet i sig är stressigt. R4 tycker inte det är stressigt generellt sett men säger att det kan bli det ibland om det är mycket arbete på gång. Även R8 förklarar att hen kan uppleva stress tillfälligt, men upplever inte arbetet som kontinuerligt stressigt.

Sex respondenter (6 av 14) upplever stressnivån i sitt arbete som varierande (R5, R6, R9, R10, R11 och R14). R5 förklarar att hen upplever att det alltid finns lite stress närvarande i sitt arbete. R6 nämner att hen upplever stressnivån som ”medelmåttlig” men inte ”jättehög”, samtidigt som den varierar till stor del.

Två respondenter (2 av 14) upplever stressnivån som ”ganska hög” (R7 och R12). R7 förklarar att stressnivån varierar mycket men att den är ganska hög ibland, men att det går att varva ner: ”Det är både och, alltså den är ganska hög emellanåt, samtidigt som man kan varva ner om vi säger så.”

Även R12 upplever stressnivån som varierande men stundvis ganska hög i samband med att det är mycket att göra (hög arbetsbelastning).

#### **5.2.5. Respondenternas tidigare erfarenheter av röststyrning**

Fyra respondenter (4 av 11) hade ingen tidigare erfarenhet av röststyrning (R1, R3, R5 och R13).

Fyra respondenter (4 av 11 respondenter som frågades) hade tidigare använt röststyrning i sin mobiltelefon (R2, R4, R8 och R10). R4, R8 och R10 sa specifikt att de använt Apples talassistent *Siri* tidigare. R8 upplevde användningen av röststyrning i sin mobil (*Siri*) som bra och smidig: ”Jo, det var smidigt. Det funka bra, röstigenkänningen tycker jag är häftigt.” R10 upplevde också användningen av *Siri* som smidig: ”Jo men det är smidigt att man bara kan säga namnet på den man vill ringa upp så ringer den automatiskt.”

Även R4 tyckte att Siri fungerade bra. Respondenten frågades även om hen upplevt några problem vid användningen av Siri och respondenten nämnde då att systemet ibland kan missuppfatta talinmatningen: ”Det är väl att den kan missuppfatta ibland, förutom det så är det bra.” (R4).

Två respondenter (2 av 11) hade tidigare erfarenhet av röststyrning i automatiserade telefonsystem (där information både tas emot och matas in via tal) (R6 och R9). R6 upplevelse av interaktionen med automatiserade telefonsystem var generellt sett ”så där”. R6 förklarade att: ”Man får ju tala väldigt tydligt för att komma rätt”. R9 uttryckte en allmänt negativ upplevelse av dessa system, vilket främst berodde på att hen upplevt att systemen missuppfattade vad hen försökte säga. Både R6 och R9 upplevde alltså problem relaterade till att systemen (till vis del) feltolkade talinmatningen.

R14 hade tidigare använt röststyrning i sin bil och upplevde att det fungerade: ”förvånansvärt bra”.

### **5.2.6. Respondenternas attityd till potentiell användning av röststyrning i sitt arbete**

Sex respondenter (6 av 11 respondenter som frågades) indikerade att de var positivt inställda till att använda röststyrning i sitt arbete ifall det skulle implementeras i någon form (R2, R6, R4, R8, R9 och R10). Följande citat från två respondenter belyser de positiva attityderna kring potentiell användning av röststyrning i deras arbete:

”Allt är bra som minimerar klick och gör arbetet mer effektivt i fabriken.” (R10).

”Inga problem, jag välkomnar nyheter.” (R8).

R2, R4 och R9 poängterar dock att om ett röstbaserat system ska användas måste det *fungera* bra, de menar att systemet måste kunna tolka talinmatningen korrekt. Samtidigt förklarar R4 att hen tror det skulle bli rätt avancerat att använda röststyrning i sitt arbete, men tror ändå att det kan finnas möjligheter.

Två respondenter (2 av 12) klassificerades (baserat på deras svar) som neutralt inställda till användningen av röststyrning i sitt arbete (R3 och R13). R3 indikerade att röststyrning skulle kunna vara användbart i en tyst miljö, men var tveksam till användningen av det i fabriken i och med ljudnivån. R13 var varken negativt eller positivt inställd, men kunde dock inte se några möjligheter själv för användning av röststyrning i sitt arbete.

Tre (3 av 11) respondenter indikerade att de var *negativt* inställda till användningen av röststyrning i sitt arbete (R1, R5, R7). R1 kunde inte se några möjligheter för användningen av röststyrning i sitt arbete och hade svårt att tro att det skulle fungera rent praktiskt: ”Jag har svårt att tänka mig att det kan funka överhuvudtaget.” R5 förklarade att hen trodde det kunde bli svårt att använda röststyrning för sina arbetsuppgifter, men att det även berodde på hur väl systemet faktiskt fungerade: ”[...] beror på hur bra röstsystemet funkar... Men det är ju svårt att röststyra vart man vill ändra”. R5 syftade på att det till exempel skulle kunna bli svårt att göra ändringar i system via röststyrning

### 5.2.7. Möjligheter för röststyrning respondenterna kunde se

Fem (5) respondenter nämner möjligheter för röststyrning associerat till inmatning (R2, R4 R6, R8 och R11), till exempel möjligheten att tala in anteckningar i systemet istället för att skriva vid exempelvis en felbeskrivning. R6 frågades om hen skulle kunna tänka sig att föra anteckningar via tal som ett alternativ till att skriva, vilket R6 visade sig vara positivt inställd till. R8 nämnde möjligheten att föra anteckningar via tal: ”Man skriver ju mycket så om man skulle kunna använda det till textinmatning så hade ju det kunnat förenkla tror jag.”

Fyra (4) respondenter svarade att de trodde det skulle kunna gå snabbare att använda röstinmatning framför att skriva (R2, R4 R6 och R11).

”Istället för att använda tangentbord går det säkert snabbare att använda rösten gissar jag.” (R6).

R2 nämner dessutom att det skulle kunna vara lättare att föra längre och mer detaljerade anteckningar vid exempelvis en arbetsbeskrivning via röstinmatning.

Fyra (4) respondenter nämnde möjligheten att interagera och ge kommandon till system via röst (R4, R6, R10 och R14).

”Ja, det skulle vara väldigt bra om man kunde ropa upp till exempel att man vill ha en sammanställning av en viss artikels utfall om något.” (R10).

”[...] det kan säkert hänga ihop med det jag sysslar med, att man begär att systemet sammanställer, och via röstkommandon gör den det utan att jag behöver använda musen och klicka in rutor och sådär.” (R6).

”[...] det skulle ju säkert vara snabbare, slipper att skriva allting. [...] att bara kunna säga spara det där och så gör den det automatiskt. Det vore ju inte fel.” (R4).

R14 nämner att operatörgränssnitten på maskiner skulle kunna vara röststyrda: ”Nej men jag ser väl operatörgränssnitten till våra bearbetningsmaskiner definitivt skulle kunna vara röststyrta.”

R5 nämnde att operatörer skulle kunna ta emot röstbaserad information via hörselkåpor. R5 menar att de till exempel skulle kunna ta emot felbeskrivningar relaterade till det arbetet de utför via röst i sina hörselkåpor: ”[...] men det finns ju system, jag vet inte om DUGA har det men det finns ju andra system för som operatören kan få med kåpor ett problem uppläst när de är på en annan maskin.”

### **5.2.8. Begränsningar för röststyrning respondenterna kunde se**

Åtta respondenter (8) kunde inte se några begränsningar eller nackdelar associerade med den potentiella användning av röststyrning i deras arbete.

Fyra respondenter (4) nämnde ljudnivån i fabriken som ett potentiellt hinder. De syftade på att bullret i fabriken skulle kunna störa ett röstbaserat system. R1 hade svårt att tro att det skulle fungera rent praktiskt: ”Jag har svårt att tänka mig att det kan funka överhuvudtaget.”

R1 förklarade även att hen trodde det skulle ta längre tid att mata in text/data via röst gentemot tangentbord vid utförandet av sina arbetsuppgifter.

R5 nämnde driftsäkerhet som ett eventuellt problem, R5 syftade på att ett röstbaserat system kanske skulle kunna störa produktionen på något sätt.

### **5.2.9. Övriga resultat från intervjuerna**

Två respondenter (R2 och R7) förklarade att om det är något krångel eller problem med en maskin brukar respondenten först ta del av information om det i datorn för att sedan gå till maskinen för att undersöka eller utföra service på maskinen. Båda respondenter nämner att de ibland även får gå tillbaka till datorn för mer information eller för att anteckna vad de utfört för service i ett system:

”Ja, det är hela tiden så, fram och tillbaka, och det man tittat på i praktiken det kanske man skriver in sen för att underlätta vid nästa gång.” (R2).

R14 nämnde sekretess- och säkerhetsaspekter associerat med röststyrning. Respondenten syftade på huruvida talinmatning från användaren till ett röstbaserat system skulle sparas eller inte och risken för avlyssning via denna typ av system.

## 6. Slutsatser och diskussion

Syftet med detta arbete har varit att översiktligt utforska och försöka skapa en bild av vilka potentiella behov det kan finnas för röstbaserade system inom fabriksmiljöer. Följande frågeställning formulerades för arbetet: *Vad finns det för generella behov relaterade till röstbaserade system inom fabriksmiljöer?*

Undersökningen genomfördes i en av Volvos fabriker i Skövde där det genom observationer och intervjuer med arbetare samlades in data som sedan analyserades och sammanställdes. I följande underkapitel följer en redogörelse och diskussion kring de slutsatser som kunnat dras baserat på resultaten från undersökningen. Det vill säga vilka behov som identifierats i formen av möjligheter och begränsningar för röstbaserade system.

### ***6.1. Identifierade behov i formen av möjligheter och begränsningar för röstbaserade system***

Det observerades att bearbetningsenheter i fabriken hade gränssnitt som lät operatörer ta del av information om enheterna samt programmera/konfigurera dem. I relation till användningsområdet *kommando och kontroll* för röststyrning skulle det potentiellt vara möjligt för operatörer att även kunna programmera eller konfigurera dessa enheter via röstkommandon. Att konfigurera dem via röstkommandon hade dessutom varit fördelaktigt ifall, precis som Nielsen (2003) nämner, operatören tillfälligt är oförmögen att interagera med systemet via sina händer. Fyra respondenter kunde dessutom se möjligheter att interagera med och kontrollera system via röstkommandon i relation till sina arbetsuppgifter. En respondent nämnde specifikt att hen trodde bearbetningsenheterna skulle kunna styras eller konfigureras via röstkommandon.

Vid observationerna noterades det även att det finns flertalet stationära datorer i fabriken som såg ut att användas för att ta del av exempelvis arbetsinstruktioner eller för att föra anteckningar, till exempel i formen av arbets- eller felbeskrivningar. Dessa antaganden bekräftades i intervjuerna där respondenterna förfrågades om de matade in eller tog emot någon form av information via systemen. En möjlighet som går att utvärdera baserat på detta är möjligheten för arbetare att föra anteckningar via diktering där talinmatningen omvandlas till text. Alternativt skulle beskrivningen kunna sparas som en ljudfil.

Totalt nämnde fem respondenter möjligheter relaterade till inmatning via röstbaserade system i sitt arbete, till exempel diktering av anteckningar. Att diktera anteckningar skulle, liknande som det nämns ovan angående röstkommandon, kunna vara användbart ifall användare är upptagna med att exempelvis utföra någon form av serviceuppgift eller liknande och för tillfället är oförmögna att skriva ner anteckningar. Med det sagt kanske det inte är nödvändigt att de för anteckningar samtidigt som de arbetar. Däremot, i relation till detta, nämnde två respondenter att de vid utförandet av serviceuppgifter ibland får gå till datorn för att anteckna vad de gjort, eller för att ta del av mer information (till exempel om det rapporterade felet), för att sedan gå tillbaka och fortsätta arbetet igen. Om till exempel ett portabelt röstbaserat dikteringssystem skulle finnas tillgängligt hade arbetarna istället kunna föra anteckningar kontinuerligt medan de arbetar och skulle därmed slippa gå till sin dator för att skriva ner anteckningar.



Fyra respondenter nämnde dessutom att de trodde det skulle kunna gå snabbare att föra anteckningar via diktering framför att skriva ner dem. En respondent nämnde vidare att det skulle vara lättare att föra längre och mer detaljerade anteckningar via diktering, till exempel om arbetet/service de utfört. Samtidigt sa en respondent att hen trodde det skulle gå långsammare att mata in text eller data via röststyrning. Värt att nämna är att dessa två respondenter utförde olika sorters arbetsuppgifter. Behoven för röstbaserade system skiljer sig uppenbarligen åt beroende på arbetsuppgifterna, vilket är en aspekt som diskuteras vidare senare i detta kapitel. De två ovannämnda respondenterna hade dessutom två olika *attityder* till användningen av röstbaserade system. En av dem hade en positiv attityd, och den andra en negativ. Attityd är också en aspekt som kan påverka behov och även denna aspekt diskuteras vidare senare i detta kapitel.

En respondent nämnde möjligheten för operatörer att få återkoppling via röst. Respondenten menade att operatörer skulle kunna få röstbaserad information via hörselkåpor medan de arbetar. En observation relaterat till detta var att många arbetare använder hörselskydd (troligtvis på grund av bullernivån i fabriken), vilket vid en första anblick skulle kunna ses som en begränsning för att ta emot återkoppling/information via röst. En annan observerad begränsning var att vissa operatörer jobbade i närhet till varandra. Om alla arbetare skulle motta röstbaserad återkoppling samtidigt från olika system skulle det troligtvis bli ansträngande att urskilja vilken information som är ämnad för vem. Men precis som respondenten nämner skulle det vara möjligt att använda hörselskydd med inbyggda högtalare, vilket möjliggör specifik återkoppling till varje arbetare utan att störa andra. Baserat på detta identifierades följande möjlighet: möjligheten att ta emot instruktioner eller återkoppling via röstbaserade system.

Den mest uppenbara eventuella begränsningen som gick att observera var ljudnivån (bullernivån) i fabriken, vilket även fyra respondenter nämnde som ett potentiellt hinder. Frågan som har uppstått är huruvida röstbaserade system skulle kunna uppfatta talinmatning korrekt i miljöer med hög ljudnivå. Irritation skulle troligtvis uppstå hos arbetare ifall systemen uppfattade inmatning felaktigt och de blev tvungna att upprepa sig flertalet gånger, vilket skulle kunna leda till en negativ användarupplevelse. Däremot visade Pires (2005) undersökning att det röstbaserade system han använde kunde uppfatta och tolka röstkommandon även i en miljö med hög ljudnivå (se kapitel 2.3.). Dock kan effektiviteten självklart skilja sig mellan olika miljöer med varierande ljudnivåer och mellan olika system. Samtidigt visade det spontana testet av talinmatning vid tre olika områden i fabriken (se kapitel 5.1.5.) att dikteringsfunktionen i författarens smartphone kunde tolka inmatningen relativt effektivt, trots bullernivån. Det är därmed möjligt att röststyrning skulle fungera inom vissa områden i fabriken, men inte i andra. Slutsatsen som går att dra från detta är att potentiella system först måste testas i den miljö som det ämnas användas inom för att se huruvida talinmatning kan uppfattas och tolkas korrekt (vilket diskuteras vidare i kapitel 6.2.).

Som nämndes i kapitel 4.2.1. frågades respondenterna om deras tidigare erfarenheter och attityd till röstbaserade system. Om användarna till exempel har tidigare negativa erfarenheter av röststyrning, eller om de har en allmänt negativ inställning till användningen av dessa system, kan det leda till motvillighet att använda dem även i framtiden. Totalt hade sju respondenter tidigare erfarenhet av röstbaserade system (sju av elva respondenter som frågades kring detta).

Tre respondenter hade mer specifikt tidigare erfarenhet av Apples talassistent Siri, vilket alla tre respondenter upplevt som positiv. En av respondenterna nämnde dock att systemet (Siri) ibland kunde missuppfatta talinmatningen.

Två andra respondenter nämnde också problem relaterade till felaktig tolkning av talinmatning vid interaktionen med automatiserade telefonsystem, och deras upplevelse av interaktionen med dessa system uppfattades därmed som negativa.

Sex respondenter (sex av elva som frågades) indikerade att de var positivt inställda till att använda röststyrning i sitt arbete. Två respondenter ansågs ha en neutral inställning och tre indikerade en negativ inställning till den potentiella användningen i sitt arbete. Fyra respondenter tryckte på att dessa system även måste *fungera* om de ska användas (till exempel kunna tolka talinmatning korrekt). Två respondenter poängterade även att det kunde bli avancerat att använda röststyrning för vissa arbetsuppgifter (till exempel att göra ändringar i system via röstkommandon). Slutsatsen som går att dra utifrån detta är att om röstbaserade system skulle utvecklas för användning inom fabriksmiljöer är det viktigt att de är enkla att använda och faktiskt är användbara, för att minimera eventuell motvillighet till att använda dem. I relation till detta är de tekniska förutsättningarna intressanta att reflektera över; är röstbaserade system tillräckligt utvecklade i nuläget för att faktiskt kunna "fungera" korrekt? Är det realistiskt och möjligt rent praktiskt att till exempel styra eller programmera enheter via röstkommandon på ett effektivt och enkelt sätt? När det till exempel kommer till tolkning av talinmatning visade exempelvis Koesters (2001) litteraturstudie att effektiviteten (relaterad till korrekt tolkning av talinmatning) för CRS-system uppmätts till 95% (se kapitel 2.4.). Däremot visade Karat et al. (2000) att det kräver mer kognitiv ansträngning samt tid att korrigera fel i text (vid exempelvis diktering) via röstkommandon i jämförelse med att göra det via tangentbord och mus. En möjlig förklaring till detta som de nämner är att röststyrda gränssnitt kanske inte är optimala när det kommer till textredigering eller att användarna i deras undersökning inte var medvetna om vilka röstkommandon som fanns tillgängliga för textredigering. Det här relaterar till något Koester (2001) nämner som en barriär för implementering av röstbaserade system; användare måste först lära sig vilka röstkommandon som är tillgängliga och hur själva interaktionen via dessa kommandon fungerar för att kunna använda systemen. Det här indikerar att röststyrning kanske inte bör vara den enda tillgängliga interaktionsformen för vissa system utan att det snarare passar bättre som ett komplement till andra interaktionsformer, i alla fall när det kommer till diktering. Huruvida den ökade tidsåtgången och kognitiva ansträngningen som Karat et al. (2000) nämner även skulle uppstå vid utförandet av andra uppgifter via röststyrning är något som skulle kunna undersökas vidare.

Sammanfattningsvis har det alltså gått att identifiera vissa möjligheter och begränsningar för röstbaserade system inom den undersökta fabriksmiljön, och resultaten från detta arbete kan eventuellt generaliseras till liknande fabriksmiljöer (baserat på miljöbeskrivningen av fabriksmiljön). Kontentan av detta är att de tekniska och praktiska möjligheterna samt begränsningarna skulle vara intressanta att undersöka vidare i mer detalj. Med andra ord: undersöka hur dessa system faktiskt fungerar i praktiken. En fortsättning på denna diskussion följer i nästa kapitel.

## **6.2. Förslag på fortsatt arbete inom ämnesområdet**

Ifall detta ämnesområde skulle undersökas vidare, antingen inom fabriken som undersöktes i detta arbete eller inom andra fabriksmiljöer, skulle det första steget förslagsvis vara att först testa effektiviteten för tolkning av talinmatning i relation till ljudnivån inom fabriken eller specifika områden inom fabriken av intresse. Ett enkelt sätt att testa detta skulle förslagsvis vara genom att använda en mobiltelefons inbyggda dikteringsfunktion i fabriksmiljön och se hur väl den tolkar talinmatningen.

Ifall det visar sig att tolkningen av talinmatning är tillräckligt effektiv skulle nästa steg förslagsvis vara att vidare undersöka och testa de praktiska och tekniska möjligheterna och begränsningarna för röstbaserade system. Skulle det exempelvis vara praktiskt att styra eller programmera bearbetningsenheter eller maskiner via röstkommandon? Det här skulle antingen kunna undersökas genom vidare litteraturstudier eller förslagsvis genom att utveckla ett simplare system som möjliggör röststyrning av enheter och därefter testa det rent praktiskt. Ifall det visar sig att det fungerar tillräckligt bra skulle steget efter det kunna vara att undersöka specifika system, enheter och arbetsuppgifter i mer detalj. Mer djupgående behovs- och uppgiftsanalyser hade kunnat genomföras för att utröna möjligheter för specifika uppgifter och interaktioner. Därefter skulle en kravanalys kunna genomföras och en kravspecifikation tas fram som skulle kunna användas av de som utvecklar systemet.

## **6.3. Etiska aspekter relaterade till ämnesområdet**

En respondent nämnde under intervjutillfället eventuella sekretess- och säkerhetsaspekter relaterade till röstbaserade system. Respondenten lyfte frågan kring hur talinmatningen som görs till systemet skulle hanteras. Skulle den till exempel kunna sparas och lagras? Skulle systemet alltid "lyssna" på vad användarna säger och vad medför detta för eventuella etiska risker? Ifall talinmatning skulle sparas av diverse skäl, till exempel för utvärdering och förbättring av systemets prestanda, medför detta såklart vissa etiska och sekretessmässiga risker som måste tas i åtanke. Till exempel hur denna data skulle hanteras och lagras, vem som skulle få tillgång till den och för vad den skulle användas. Troligtvis skulle även informerat samtycke behövas från de som skulle använda systemen. Det här behöver inte nödvändigtvis vara något större hinder. Om enbart talinmatning i form av exempelvis kommandon lagras; hur känslig skulle denna informationen vara? Det är en aspekt som självklart behöver avvägas ifall denna typ av system skulle utvecklas och implementeras. Alternativt kanske det inte ens är nödvändigt att systemen spelar in och lagrar talinmatningen (eller annan relaterad information). Den andra farhågan, att systemet skulle kunna användas för avlyssning kan medföra ytterligare etiska risker som skulle behöva diskuteras kring och avvägas. Däremot är det svårt att se att denna typ av funktion överhuvudtaget skulle implementeras i ett system i och med att det är högst kontroversiellt. Skulle det till att börja med ens vara lagligt att använda denna typ av system för avlyssning, och skulle företaget bakom systemet ens ha något intresse av att avlyssna sina anställda? De här aspekterna och riskerna är såklart relevanta och viktiga att diskutera. Det är inte omöjligt att avlyssning skulle kunna förekomma via dessa system, men en fortsatt diskussion kring detta hör hemma i ett annat forum.

En annan intressant aspekt att diskutera kring är *vad* syftet med utvecklingen av denna typ av system skulle vara. Skulle syftet vara att skapa en bättre användarupplevelse för användarna (arbetarna), eller skulle det vara att effektivisera arbetsuppgifterna (vilket troligtvis skulle leda till ökad produktivitet och ekonomiska fördelar). Kanske skulle syftet vara både och, men innebär ökad effektivitet även att användarupplevelsen förbättras? Det är möjligt att utvecklingen och implementeringen av denna typ av system förbättrar effektiviteten vid både interaktionen med system och utförandet av arbetsuppgifterna, men användarnas upplevelse kanske samtidigt påverkas negativt. Alternativt kanske användarnas upplevelse förbättras, men deras arbetseffektivitet försämras (till exempel kanske det tar längre tid att utföra uppgifter via röststyrning). Det går att reflektera kring detta men det är svårt att dra några slutsatser utan att faktiskt utveckla någon form utav prototyp och faktiskt testa vilka potentiella effekter användningen kan ha på exempelvis användarnas upplevelse och effektiviteten. Det skulle därmed behöva göras ytterligare en avvägning inför utvecklingen mellan eventuella sociala kontra ekonomiska hållbarhetsaspekter samt konsekvenser som går att se associerade med användningen av denna typ av system.

#### **6.4. Arbetets bidrag till UXD-fältet**

Det här arbetets huvudsakliga bidrag till UXD-fältet är en inblick i potentiella behov för röstbaserade gränssnitt inom fabriksmiljöer, samt användarnas attityd till denna typ av system. Andra intressenter skulle kunna använda sig av resultatet som en grund eller inspiration för fortsatta undersökningar och empiriskt arbete inom ämnesområdet. Till exempel skulle en UX-designer, verksam inom ett projekt där röstbaserade system utforskas och ämnas utvecklas, eventuellt kunna generalisera och använda resultatet från detta arbete. Resultatet skulle i detta fall exempelvis kunna användas som underlag vid utvecklingen av ett röstbaserat system och bidra till att dessa behov tas i åtanke under utvecklingen. Huruvida det går att generalisera resultatet skulle förslagsvis kunna avgöras baserat på beskrivningen av kontexten, det vill säga fabriksmiljön, som presenteras i denna rapport. Om fabriksmiljöerna i andra eventuellt framtida projekt delar egenskaper med fabriken som undersöktes i detta arbete kan det vara möjligt att resultatet går att generalisera och utgå från som en grund.

#### **6.5. Utmaningar och lärdomar**

Som Wurhofer et al. (2015) och Obrist et al. (2011) nämner finns det vissa utmaningar associerade med att genomföra undersökningar inom fabriker. Dessa svårigheter menar de främst har med att undersökningar inte får komma i vägen för eller störa arbetet som utförs inom fabriker. Obrist et al. (2011) nämner också svårigheter associerade med tidsbegränsningar för arbetare inom fabriker som begränsar möjligheten att involvera dem. Dessa utmaningar gjorde sig även kända i detta arbete. Undersökningarna fick inte störa det pågående arbetet som genomfördes inom fabriken. Begränsningarna detta medförde praktiskt för observationsstudierna var att författaren av denna rapport inte fick störa arbetarna och ställa en mängd frågor medan de utförde sina arbetsuppgifter. Författaren inser att det hade varit fördelaktigt att genomföra mer deltagande etnografiska observationer. Det vill säga observationer där författaren skulle anta en form av lärlingsroll, med syftet att få en mer djupgående förståelse för arbetsuppgifterna och interaktionerna med systemen inom fabriken.

I relation till detta hade det i så fall även varit fördelaktigt att avgränsa arbetet ytterligare, och fokusera behovsanalysen på exempelvis ett fåtal arbetsuppgifter, system och interaktioner med systemen. Det hade därmed varit möjligt att undersöka och utröna potentiella behov mer djupgående för dessa specifika uppgifter och interaktioner genom exempelvis en uppgiftsanalys. Utmaningarna associerade med intervjuerna hade med tidsbegränsning att göra. I och med att intervjuerna skulle genomföras med arbetarna inom fabriken på deras arbetstid var det inte möjligt att ta upp för mycket av deras tid. Författaren och handledaren på Volvo kom överens om att intervjuerna lämpligtvis skulle kunna ta fem till tio minuter per respondent. Författaren insåg även vid intervjutillfällena att de inte kunnat vara för långdragna; många respondenter undrade exempelvis hur lång tid intervjun skulle ta och förklarade att den inte fick ta *för* lång tid eftersom de hade saker att göra. Författaren önskar att intervjuerna hade kunnat vara längre och att mer sonderande frågor hade ställts eftersom det skulle kunna leda till en mer djupgående förståelse för, och insikt, i exempelvis respondenternas arbetsuppgifter. Alternativt hade det även varit intressant att genomföra någon form utav fokusgrupp som låtit arbetare inom fabriken diskutera kring ämnesområdet tillsammans, vilket hade kunnat leda till ytterligare insikter.

Om det, som nämns ovan, hade varit möjligt att genomföra dessa två metoder mer ingående, och om omfattningen på undersökningen avgränsats ytterligare, hade resultaten kunnat användas för att genomföra en mer djupgående uppgiftsanalys. Genomförandet av en sådan uppgiftsanalys hade kunnat resultera i mer specifika och konkreta behov, och resultatet från denna uppgiftsanalys hade dessutom potentiellt kunnat användas för att genomföra en kravanalys och skapa en kravspecifikation för ett röstbaserat system. Som nämns i kapitel 6.2. hade detta varit ett intressant tillvägagångsätt (om möjligt) för fortsatt arbete inom detta ämnesområde.

Vid intervjuerna ställdes frågor kring respondenternas arbetsuppgifter och deras systemanvändning. De här frågorna ställdes med syftet att försöka utröna tillfällena eller steg i utförandet av arbetsuppgifter eller interaktionerna med system där röststyrning eventuellt skulle kunna användas. Respondenterna frågades även om deras upplevelse associerad med systemanvändningen, samt hur de upplevde arbetsbelastningen och stressen i sitt arbete. Syftet med dessa frågor var att undersöka eventuella frustrationer och problem som användarna upplevde, vilket Kujala (2006) som sagt förklarar är en möjlig källa när det kommer till att undersöka och utröna behov. Dock visade det sig vid analysen av insamlad data från intervjuerna att svaren på dessa frågor var för översiktliga och det var därmed inte möjligt att utröna några direkta behov ur dem. Återigen, hade längre och mer sonderande intervjuer och kompletterande observationer varit passande, men som Obrist et al. (2011) förklarar måste undersökningar anpassas efter omständigheterna.

Det hade dock som sagt varit intressant att undersöka ämnet mer djupgående, vilket är det främsta förslaget för fortsatt arbete: en mer djupgående behovsanalys avgränsad till ett fåtal arbetsuppgifter och systeminteraktioner.

## 6.6. Sammanfattning av slutsatser

Avslutningsvis presenteras här en kortfattad sammanfattning av de slutsatser som dragits baserat på resultatet från det arbete som genomförts.

Syftet med detta arbete var att undersöka behov för röstbaserade system inom fabriksmiljöer i formen av begränsningar och möjligheter.

De huvudsakliga identifierade möjligheterna är följande:

- Möjligheten för operatörer att programmera/konfigurera bearbetningsenheter via röststyrning (vilket relaterar till användningsområdet *kommando och kontroll* för röstbaserade system).
- Möjligheten att föra anteckningar via diktering, till exempel anteckningar i formen av arbets- eller felbeskrivningar.
- Möjligheten att få återkoppling/information via röst om exempelvis rapporterade fel och service som behöver utföras.

Följande huvudsakliga begränsningar för röstbaserade system identifierades:

- Den huvudsakliga begränsningen berör ljudnivån (bullernivån) inom fabriksmiljöer, som dock kan variera mellan olika fabriksmiljöer och områden inom fabriker.
- Tekniska utmaningar och begränsningar; huruvida det är tekniskt möjligt och praktiskt att faktiskt använda röstbaserade system för specifika uppgifter är en aspekt som skulle vara intressant att undersöka vidare.

Ytterligare en påverkande aspekt är användarnas tidigare erfarenhet och attityd till röstbaserade system. Om användarna har negativa erfarenheter eller en allmänt negativ attityd till att använda denna typ av system medför det vissa begränsningar. Resultatet från undersökningen i detta arbete indikerade att respondenternas tidigare erfarenhet och attityd till röstbaserade system var blandad, men övervägande positiv.

Som det diskuterades i föregående kapitel finns det vissa utmaningar associerade med att genomföra undersökningar i fabriksmiljöer. Dessa utmaningar relaterar främst till att undersökningar inte får störa det pågående arbetet inom fabriksmiljöer, vilket även satte begränsningar för metoderna som ämnades användas i detta arbete. Till följd av dessa utmaningar var undersökningen tvungen att anpassas efter de rådande omständigheterna. Dessa anpassningar kanske inte alltid var fördelaktiga, men samtidigt medförde dessa utmaningar en rad lärdomar kring utförandet av empiriskt arbete för författaren av denna rapport.

## Referenser

- Ayres, T., & Nolan, B. (2006). Voice activated command and control with speech recognition over WiFi. *Science of Computer Programming*, 59(1-2), 109-126. DOI: 10.1016/j.scico.2005.07.007
- Berndtsson, M., Hansson, J., Olsson, B., & Lundell, B. (2008). *Thesis Projects: A Guide for Students in Computer Science and Information Systems* (Vol. 2). London: Springer. ISBN: 13: 978-1-84800-009-4
- Barnum, C. M. (2011). *Usability Testing Essentials: Ready, Set... Test!* Burlington, MA: Morgan Kaufmann.
- Cooper, A., Reinmann, R., Cronin, D., & Noessel, C. (2014). *About Face: The Essentials of Interaction Design* (Vol. 4). Indianapolis: John Wiley & Sons, Inc. ISBN: 978-1-118-76657-6.
- Hassenzahl, M., & Tractinsky, N. (2006). User experience. *Behaviour & Information Technology*, 25(2), 91-97. DOI: <https://doi.org/10.1080/01449290500330331>
- Jazdi, N. (2014). Cyber Physical Systems in the Context of Industry 4.0. *IEEE International Conference on Automation, Quality and Testing, Robotics* (ss. 1-4). Cluj-Napoca: IEEE. DOI: 10.1109/AQTR.2014.6857843
- Kambeyanda, D., Singer, L., & Cronk, S. (1997). Potential Problems Associated with Use of Speech. *Assistive Technology*, 9(2), 95-101. DOI: <https://doi.org/10.1080/10400435.1997.10132301>
- Karat, J., Horn, D. B., Halverson, C. A., & Karat, C. M. (2000). Overcoming Unusability: Developing efficient strategies in speech recognition systems. *CHI '00 Extended Abstracts on Human Factors in Computing Systems* (ss. 141-142). New York: ACM. DOI: <https://doi.org/10.1145/633292.633372>
- Koester, H. H. (2001). User Performance With Speech Recognition: A Literature Review. *Assistive Technology*, 13(2), 116-130. DOI: <https://doi.org/10.1080/10400435.2001.10132042>
- Kujala, S. (2006). Effective user involvement in product development by improving the analysis of user needs. *Behaviour & Information Technology*, 27(6), 457-473. DOI: 10.1080/01449290601111051
- Kulyukin, V. (2004). Human-Robot Interaction Through Gesture-Free Spoken Dialogue. *Autonomous Robots*, 16(3), 239-257. DOI: <https://doi.org/10.1023/B:AURO.0000025789.33843.6d>
- Lowgren, J. (u.å.). *Interaction Design - brief intro | The Encyclopedia of Human-Computer Interaction, 2nd Ed.* Hämtat från Interaction Design Foundation: <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/interaction-design-brief-intro> [2019-02-12]
- Maguire M. & Bevan N. (2002) User Requirements Analysis. In: *Hammond J., Gross T., Wesson J. (eds) Usability. IFIP WCC TC13 2002. IFIP — The International Federation for Information Processing*, vol 99. Springer, Boston, MA. DOI: [https://doi.org/10.1007/978-0-387-35610-5\\_9](https://doi.org/10.1007/978-0-387-35610-5_9)

- Nass, C., & Brave, S. (2005). *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*. London: MIT Press. ISBN: 9780262640657.
- Nielsen, J. (2003). *Voice Interfaces: Assessing the Potential*. Hämtat från Nielsen Norman Group: <https://www.nngroup.com/articles/voice-interfaces-assessing-the-potential/> [2019-01-30]
- Obrist, M., Reitberger, W., Wurhofer, D., Förster, F., & Tscheligi, M. (2011). User Experience Research in the Semiconductor Factory: A Contradiction? i P. Campos, N. Graham, J. Jorge, N. Nunes, P. Palanque, & M. Winckler, *Human Computer Interaction - INTERACT 2011* (Vol. 6949, ss. 144-151). Berlin, Heidelberg: Springer. DOI: [https://doi.org/10.1007/978-3-642-23768-3\\_12](https://doi.org/10.1007/978-3-642-23768-3_12)
- Patton, M. Q. (2015). *Qualitative Research & Evaluation Methods: Integrating Theory and Practice. (4:e uppl.)*. Thousand Oaks, CA: SAGE Publications. ISBN 9781412972123.
- Pires, N. J. (2005). Robot-by-voice: Experiments on commanding an industrial robot using the human voice. *Industrial Robot: An International Journal*, 32, 505-511. DOI: <https://doi.org/10.1108/01439910510629244>
- Preece, J., Rogers, Y., & Sharp, H. (2015/2016). *Interaktionsdesign: Bortom Människa-Dator-Interaktion*. (L. Svensson, & M. Drangel, Övers.) Lund: Studentlitteratur.
- Reul, M. (2009). Bringing usability to industrial control systems. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems (CHI EA '09)*. ACM, New York, NY, USA, 3335-3340. DOI: 10.1145/1520340.1520482
- Rogowski, A. (2012). Industrially oriented voice control system. *Robotics and Computer-Integrated Manufacturing*, 28(3), 303-315. DOI: <https://doi.org/10.1016/j.rcim.2011.09.010>
- SS-EN ISO 9241-210. (2010). *Ergonomi vid människa-systeminteraktion - Del 210: Användarcentrerad design för interaktiva system*. Stockholm: SIS förlag.
- Unger, R. & Chandler, C. (2012). *A Project Guide to UX Design: For user experience designers in the field or in the making (2:a uppl.)*. Berkeley, CA: New Riders. ISBN: 9780321815386
- Wikberg Nilsson, Å., Ericson, Å. & Törlind, P. (2015). *Design Process och Metod*. Studentlitteratur. ISBN: 9789144108858.
- Wurhofer, D., Fuchsberger, V., Meneweger, T., Moser, C., & Tscheligi, M. (2015). Insights from User Experience Research in the Factory: What to Consider in Interaction Design. i J. Abdel Nocera, B. Barricelli, A. Lopes, P. Campos, & T. Clemmensen, *Human Work Interaction Design. Work Analysis and Interaction Design Methods for Pervasive and Smart Workplaces*. (Vol. 468, ss. 39-56). Cham: Springer. DOI: [https://doi.org/10.1007/978-3-319-27048-7\\_3](https://doi.org/10.1007/978-3-319-27048-7_3)



# Bilagor

## Bilaga 1 - Intervjufrågor och information till respondenter

### Introduktion/bemötande

*"Hej! Skulle du vilja ha lite fika och svara på några frågor för en undersökning vi genomför här?"*

*"Vi är två studenter från Högskolan här i Skövde och vi gör just nu vårt examensarbete i samarbete med Volvo där vi undersöker några aspekter relaterade till arbetsuppgifter och datorsystemen i fabriken."*

*"Innan vi börjar, är det OK att vi spelar in intervjun i syfte att underlätta vid analysen av svaren?"* **Efter svar angående inspelningen:** *"Inspelningarna kommer enbart hanteras av oss och eventuellt våra handledare på högskolan. Efter arbetet är slutfört kommer inspelningarna även att raderas."*

*"Även dina svar på frågorna kommer endast behandlas utav oss, och eventuellt våra handledare på högskolan. Resultaten från intervjuerna kommer presenteras i vår rapport som är en offentlig handling, men dina svar kommer vara anonyma och inte kunna kopplas till dig personligen. Frågorna är helt frivilliga, om du inte vill svara på någon specifik fråga så behöver du inte det, och du kan när som helst avbryta intervjun."*

---

### Del 1. - Demografiska uppgifter

*"Först tänkte jag ställa ett par frågor om dig själv och därefter följer några frågor om ditt arbete och hur du upplever det."*

1. **Vad är din ålder?**
  2. **Hur länge har du arbetat i fabriken?**
- 

### Del 2. - Arbetsuppgifter och användarupplevelse

3. **Vad är det du huvudsakligen gör i fabriken?**
4. **Kan du beskriva en specifik arbetsuppgift du utför? (Om dem inte redan gjort det i förra frågan)**
5. **Hur upplever du arbetsbelastningen i ditt arbete?**
6. **Hur upplever du stressnivån i ditt arbete?**
  - 6.1. **Om de upplever någon form av stress:** Kan du beskriva vad du gör när du upplever denna stress?

### Del 3. - Användning av datorsystem

*"Nu tänkte jag gå vidare och ställa några frågor angående användningen av datorsystem i fabriken och ditt arbete."*

#### 7. Använder du något datorsystem i arbetet och till vad i så fall?

7.1. **Om ja:** Kan du beskriva hur du använder systemet till någon arbetsuppgift?

7.2. Tar du emot någon form av information via systemet? **Om ja:** Kan du beskriva detta närmare?

7.3. Förs det anteckningar eller matas in information av någon form i systemet?  
**Om ja:** Hur går du tillväga då?

#### 8. Hur upplever du det är att använda systemet/en?

#### 9. Har du upplevt några problem med systemet/en?

---

### Del 4. - Röststyrning

*"Då var det dags för några avslutande frågor relaterade till röstbaserade datorsystem."*

*"Med röstbaserade system syftar jag på system som det på diverse sätt går att använda genom tal. Det kan exempelvis röra sig om att ge kommandon eller få information från ett system via tal, med mera."*

#### 10. Har du någon tidigare erfarenhet av röstbaserade system?

10.1. **Om ja:** be dem beskriva användningen.

10.2. **Om ja:** hur upplevde du användningen?

#### 11. Hur tror du det skulle vara att använda ett röstbaserat system vid utförandet av någon av dina arbetsuppgifter?

11.1. Kan du se några möjligheter för röststyrning i ditt arbete? Vad i så fall?

11.2. Kan du se några nackdelar? Vad i så fall?

---

### Avslutning

*"Det var allt vi hade, har du något du vill tilläga eller några frågor till oss?"*

*"Då tackar vi för din tid och medverkan och önskar dig en fortsatt trevlig dag."*