

Safe Speech by Knowledge

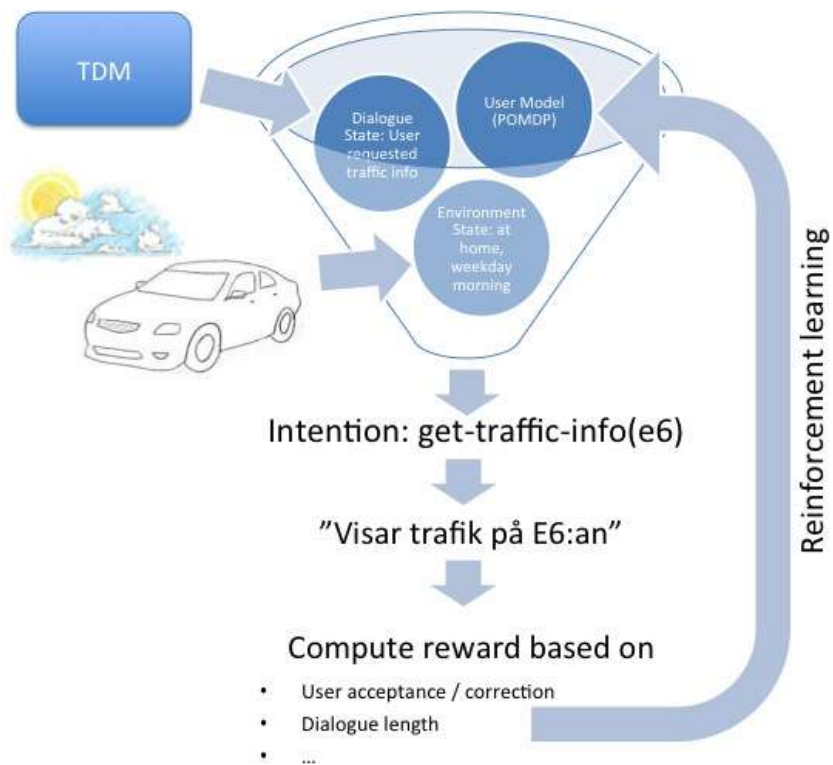


Table of Contents

1	Sammanfattning	<u>23</u>
2	Bakgrund	3
3	Syfte	4
4	Genomförande	4
4.1	Analys	<u>45</u>
4.2	Användningsstudie	5
4.3	Implementation	6
4.3.1	Dialogfunktionalitet	6
4.3.2	Rapid Prototyping-plattform	6
5	Resultat	<u>67</u>
5.1	Demonstrator/Implementation	<u>67</u>
5.2	Kunskap om användarupplevelse kring prediktiv dialog	<u>89</u>
5.3	TDM på embeddedplattform	9
5.4	Utvecklingsbeskrivning VCC	<u>910</u>
5.5	Bidrag till FFI-mål	11
5.5.1	Implementera ett lärande system som reducerar mängden repetitiva och informationsfattiga yttranden	11
5.5.2	Bygga UX-kunskap kring prediktiv dialog	11
5.5.3	TDM på embeddedplattform	11
5.5.4	Kunskapsbyggande om att ta in ny funktionalitet från mobilplattform in i bilen.	11
5.5.5	Patentansökan	<u>1112</u>
6	Spridning och publicering	12
6.1	Kunskaps- och resultatspridning	12
6.2	Publikationer	12
7	Slutsatser och fortsatt forskning	12
8	Deltagande parter och kontaktpersoner	13

1 Sammanfattning

Dagens förarmiljö har blivit alltmer komplex och ställer höga krav på interaktionsdesign. Dessutom utförs i allt större utsträckning sekundäruppgifter av förare via smartphones och annan teknologi som bidrar till distraktion och därigenom ökar risken för olyckor. Att

designa interaktion där föraren blir mindre distraherad, och kan hålla händerna på ratten och ögonen på vägen, blir därför en viktig uppgift. I detta projekt är syftet att designa, implementera och testa ett multimodalt adaptivt dialogsystem i syfte att reducera interaktionstid för att minska distraktion och därmed öka säkerheten i förarmiljö. Projektet syftar också till att bygga kunskap om hur prediktiva modeller kan användas för att designa talad interaktion där hög användarupplevelse kan uppnås. Projektet har utförts i tre steg: analys, användningsstudie och implementation och resulterat i en fungerande implementation som reducerar dialogen baserat på användarens körbeteende. Systemet minskar också andelen repetitiva frågor från systemet vilket enligt projektets användningsstudie dels reducerar interaktionstiden, dels bidrar till en positiv användarupplevelse för föraren. Vidare har användningsstudien givit upphov till en uppsättning användarcentrerade designprinciper för adaptiv talad interaktion som kan användas för framtida utvecklingsprojekt. Slutligen har en användningstestmetod optimerad för snabb prototypning av talad interaktion i bil tagits fram under projektets gång.

2 Bakgrund

Förardistraktion är en vanlig anledning till olyckor, och orsakas ofta av att föraren interagerar med någon sorts teknologi, t.ex. mobiltelefon, navigator eller mediaspelare. En studie, populärt kallad ”100 car study” (Neale et al., 2005) uppskattar att distraktion som orsakas av att utföra en sekundäruppgift är den största orsaken till ouppmärksamhet, och att den vanligaste sekundäruppgiften är att hantera en trådlös enhet av något slag. NHTSA (U.S. National Highway Traffic Safety Administration) uppskattar att 16 % av de dödliga olyckorna i USA orsakades av distraktion, liksom 20 % av de olyckor som ledde till allvarliga skador.

I fordonsindustrin talar man om HMI (Human-Machine Interaction) när man pratar om ett användargränssnitt. Ett sådant HMI kan eventuellt inkludera röst som input och output. I forskningsvärlden pratar man om multi-modala dialogsystem (MDS). Ett multimodalt dialogsystem möjliggör talad kommunikation mellan användare och maskiner, men kompletterar den verbala interaktionen med traditionella HMI-komponenter för interaktion såsom bildskärmar, knappar, pekskärmar etc.

TDM är Talkamatics produkt för dialoghantering. En dialoghanterare är den centrala komponenten i ett dialogsystem. TDM har ett antal attraktiva egenskaper – snabb utveckling, fri dialog samt inbyggd multimodalitet. Aktuellt fokus för vidareutvecklingen är låg distraktion och snabbare utveckling.

State-of-the-art-dialogsystem brukar inte inkludera någon aktiv användarmodellering, utan hanterar alla användare på samma sätt. Apples Siri, som kommer att introduceras i fordon, har viss modellering av användaren på så sätt att de data som användaren har angivit används av Siri. Till exempel använder Siri per default användarens hemadress

när denne säger att hon eller han vill boka en taxi. Detta utgör en statisk användarmodell, av den anledningen att den inte uppdateras av interaktionen med systemet. Google Now använder en dynamisk användarmodell (uppdateras av interaktionen med systemet), men saknar möjligheten till rik röstinteraktion.

3 Syfte

Det primära syftet med detta projekt har varit att bygga ett multimodalt dialogsystem med en dynamisk användarmodell, där användarmodellen påverkar dialogen på så sätt att längden på interaktionen som behövs för att utföra en viss uppgift minimeras: Användaren behöver inte svara på repetitiva frågor där systemet kan lista ut svaret. Genom att minimera längden på interaktionen minimeras också tiden som användaren är distraherad.

Ett sekundärt syfte har varit att vinna kunskap om prediktiva system och talad interaktion med sådana system. Hur skall de designas? Vad förväntar sig användaren? Uppstår det problem i samband med tekniken och användningen? Frågor som väckts är t ex: känner användaren sig övervakad och kartlagd när systemet i bilen vet detaljer om ens liv som man inte explicit har berättat om? Hur designar och prototyper man adaptiva talade system i förarmiljö?

4 Genomförande

Genomförandet av projektet hade tre komponenter: Analys, Användningsstudie och Implementation. Analysdelen gick ut på att skapa en grundläggande förståelse för var tekniken som utvecklas gör bäst nytta: Vilka typer av mönster är lättast att lära in, och vilka typer av applikationer är länkade till dessa data? Användningsstudien gick ut på att svara på frågor om upplevelse och interaktion vid användande: Användarens förväntan på prediktiva system, hur de skall utformas, användbarhetsproblem och –vinster, etc. Dessa komponenters resultat gick in i implementationsdelen, vilken hade som mål att bygga det faktiska systemet med användarmodell, inlärningsförmåga, stöd för resonemang och dialog om osäker kunskap samt applikationer.

4.1 Analys

Analysarbetet bestod av en litteraturstudie om adaptiva system, naturligt-språk-interaktion i förarmiljö, samt förar/användarupplevelse. Detta arbete gav dels direkt stöd i

designval, dels lade det grunden för hur empiriinsamlingen skraddarsyddes för projektet (se Användarstudie, nedan).

I analysarbetet ingick även en genomgång av data om körvanor, liksom konsumentbeteende avseende vilka teknologiska plattformar och tjänster som används i olika målgrupper (såsom smartphoneanvändning för olika åldersgrupper). Denna typ av data påverkar designval för det ekosystem som den slutgiltiga tjänsten ska ingå i.

I analysfasen ingick två omgångar workshops på plats hos VMCC. Första workshopen hade deltagare från TAB och VMCC, och fokuserade på hur upplevelsen av det prediktiva systemet skulle förenas med VCCs kärnvärden, och hur man skulle använda tekniken för att minimera användarens ”visual struggle” med att hantera tekniken. Workshopen utmynnade i ett antal förslag på use cases i enlighet med detta. Saker som betonades var att användaren inte ska känna sig övervakad av systemet. Veckan därpå gick andra workshopen av stapeln då Högskolan i Halmstad och VMCC fokuserade på användarupplevelse genom persona- och scenarioskapande. Scenariot som etablerades innehöll två primära personas (som representerar arketyper av användare av tjänsten).

Under dessa workshops utvecklades målet med att bara reducera antalet yttranden. Det ifrågasattes om att enbart reducera just antalet yttranden alltid är en förbättring. Slutsatsen av diskussionen var att en dialog som befinner sig på rätt nivå är viktigare än just en kort dialog, och att en dialog på rätt nivå ligger i området mellan uttråkning och stress. Detta innebär att upprepning av samma information skall undvikas till förmån för att användaren accepterar sådan information i form av systemförslag. Det innebär också att systemet bör ha mycket goda skäl att ta bort yttranden – eftersom borttagande av yttranden på ett oförklarligt sätt kan orsaka stress.

4.2 Användningsstudie

HH genomförde en kvalitativ användningsstudie där en prototyp av tjänsten som skissats ut i scenariona från Workshop 2 testades. Sju respondenter deltog i studien, där de fick köra en bil riggad med en prototyp av en adaptiv navigeringstjänst. Scenariot bestod av en progression i tre delar, där systemet lärde sig användarens körbeteende och anpassade dialogen därefter. Studien genomfördes som en Wizard-of-Oz-studie, där användaren var omedveten om att en människa (”trollkarlen”) skötte interaktionen. ”Trollkarlen” som sköter systemets dialogturer följer ett i förväg definierat interaktionsprotokoll baserat på teori om adaptiv interaktion, best-practice inom talad interaktion, samt scenariona som togs fram vid Workshop 2. Baserat på interaktionsprotokollet och användarens input sänder ”trollkarlen” skraddarsydda textsträngar till bilen som läses upp via en state-of-the-art text-to-speech-modul. Mellan scenariona anpassades interaktionsprotokollet för att simulera anpassning till användarens körbeteende och preferenser.

Efter varje delscenario genomfördes en semistrukturerad intervju med användaren om hennes upplevelse. Syftet var att få användaren att diskutera olika aspekter om användbarhet, användarupplevelse och upplevd stress, säkerhet, kognitiv belastning, och emotionell respons från interaktionen med systemet.

4.3 Implementation

4.3.1 Dialogfunktionalitet

Implementationen av dialogfunktionaliteten drevs som ett agilt mjukvaruprojekt, och genomfördes helt och hållet av Talkamatic. Både VCC och Högskolan i Halmstad bidrog dock med regelbunden feedback på systemet, och delade med sig av insikter och erfarenheter, i Högskolan i Halmstads fall framför allt insamlade under användarstudien.

Som alla agila projekt präglades implementationen av snabba releaser, korta iterationer, fokus på användarbehov (via HH) och testdriven utveckling. Under implementationsprojektet gjordes fem releaser:

- Release 1 – Grundläggande funktionalitet: användarmodell med inläring från interaktion. Mockad trafikinformationsapplikation.
- Release 2 – förfinad applikation, bättre data. Feedback från systemet på osäker data.
- Release 3 – verktyg för att simulera användarmodeller med grafiskt gränssnitt. Bättre feedback på osäker data.
- Release 4 – Bättre data, supportar use case från persona. Möjlighet att visualisera användarmodell. Ytterligare förbättrad feedback på osäker data.
- Release 5 – Hela systemet kör på ett *BD-SL-i.MX6*-kort.

4.3.2 Rapid Prototyping-plattform

Parallellt med utvecklingen av dialogfunktionaliteten bedrev VCC utvecklingsarbete på ett antal punkter: Dels togs en generell lösningsarkitektur fram för den typ av röstinteraktion som projektet behandlar, dels vidareutvecklades rapid-prototyping-plattformen för användning på IMX-6-kortet. Dessutom togs en prototypapplikation för användning med Talkamatics plattform fram.

5 Resultat

5.1 Demonstrator/Implementation

Det finns en demonstrator som visar interaktion med ett lärande dialogsystem, som till exempel klarar av att hantera exempeldialogen från ansökan:

```
User: Traffic information
Car: Ok. What road?
User: E6.
Car: Showing traffic on the E6
```

This is repeated on a number of occasions, and one morning...

```
User: Traffic information
Car: Showing traffic on the E6
```

Dialogsystemet lär sig från användarens interaktion med systemet och från användarens faktiska körbeteende. Händelser klassificeras efter tidpunkt på dagen, veckodag och position.

Systemet minskar behovet av repetitiva och informationsfattiga yttranden från användaren. Även innan systemet slutar ställa frågan, blir dialogen bättre. När systemet har börjat känna igen ett mönster, börjar den föreslå det alternativ som verkar mest troligt. Till en början sker det genom att de troligaste alternativen föreslås som första alternativ i en lista. Alternativen är dessutom markerade i avvikande färg för att göra det enklare för användaren att se dessa:

```
User: Traffic information
Car: Ok. What road?
Displaying:  [E6]
              [E45]
              [E20]
              [155]
              [40]

User: E6.
Car: Showing traffic on the E6
```

Efter ytterligare användning har systemet identifierat ett tillräckligt tydligt mönster för att ge ett förslag.

```
User: Traffic information
Car: E6, is that right?
User: Yes.
Car: Showing traffic on the E6
```

Efter ytterligare stöd för systemets hypotes börjar systemet istället informera om att ett antagande har gjorts. Om användaren är nöjd med antagandet behöver denne inte göra något, men kan bekräfta det:

```
User: Traffic information
Car: I assume E6.
User: [silence]
Car: Showing traffic on the E6
```

Användaren har alltid möjlighet att förkasta systemets förslag.

User: Traffic information
Car: I assume E6.
User: No, E45.
Car: Showing traffic on the E45

Om användaren förkastar systemets förslag utan att ge ett alternativt svar visar systemet en meny, där de mest troliga valen är sorterade överst och markerade i en avvikande färg.

User: Traffic information
Car: I assume E6.
User: No.
Car: What road?

Displaying:

[E6]
[E45]
[E20]
[155]
[40]

Då systemet är helt säkert på sin hypotes, ger systemet den önskade informationen utan att ens be användaren om parametrar:

User: Traffic information
Car: Showing traffic on the E6

Detta beteende gör sammantaget att den kognitiva belastningen och distraktionen för användaren i de allra flesta fall bör minska.

5.2 Kunskap om användarupplevelse kring prediktiv dialog

En förars huvudsakliga uppgift är att framföra sitt fordon på ett säkert sätt i trafiken (primär interaktion). I dagens förarmiljö finns dessutom sekundär interaktion i form av bland annat infotainmentsystem, klimatkontroll och navigering. Inom människa-maskin-interaktion i förarmiljö arbetas det mycket med att säkerställa att den sekundära interaktionen inte påverkar trafiksäkerheten negativt. Genom att studera användande och brukarsituation är det dock lätt att se att det även finns tertiär¹ interaktion i förarmiljö, i

¹ I viss litteratur slås sekundär och tertiär interaktion samman och benämns sekundär interaktion. I detta projekt är det dock meningsfullt att skilja på sekundär interaktion med

form av bl a smartphone-användning för exempelvis navigering eller kommunikation i form av sms eller samtal. Dessa tre nivåer av interaktion ställer allt högre krav på förarens kognitiva resurser, vilket kan påverka förarens förmåga att lösa sin primära uppgift negativt. Frånsett säkerhetsaspekterna, påverkar dessutom denna interaktionskomplexitet användarupplevelsen (user experience, UX) negativt. I detta projekt studerar vi aspekter av UX i användningen av prediktiv dialog för att (a) dels inte belasta visuella-manuella interaktionsmodliteter för att frigöra ögon händer för den primära interaktionen, (b) dels för att förkorta sekundär och tertiär interaktion så att föraren kan frigöra kognitiva resurser för den primära uppgiften.

Genom att prototypa och testa dialogsystemlösningar (se sektion 4.2) med adaptiva och prediktiva kvaliteter på representativa slutanvändare i autentisk förarmiljö har vi byggt upp kunskap kring upplevelsebaserade kvaliteter gällande talad interaktion i förarmiljö. Studiens analys verifierade positiva egenskaper och gav indikationer på problem gällande i huvudsak sex aspekter av interaktion: antropomorfism, tillit till systemet, förutsägbarhet, förståelse, integritet och kontroll. I studiens debrief-intervjuer diskuterades lösningar på systemets tillkortakommanden vilket mynnat ut i en modell för UX-orienterad design av talad adaptiv interaktion i förarmiljö. (Manuskript som rapporterar detta är vid dagens datum under bearbetning.)

Ytterligare resultat inom UX-kunskap är den metodutveckling som framtagits under projektet, som adresserar problemet med att snabbt och ekonomiskt testa funktionalitet och upplevelse i en autentisk brukarsituation. Metoden är baserad på Wizard-of-Oz-metodiken och förstärkt med en specifik verktygs- och miljöuppsättning skraddarsydd för att prototypa och testa adaptiv dialoginteraktion i bil.

5.3 TDM på embeddedplattform

Som en del i projektet har TDM flyttats över till en embedded-plattform, ett kort från Boundary Devices (*BD-SL-i.MX6*) som följer Freescales Light Sabre-specifikation. Kortet är i termer av processorkraft, minne, konnektivitet etc. representativt för de plattformar som används i dagens infotainment-plattformar. Detta innebär att vi nu har verifierat att TDM kan köras på dagens infotainmentplattformar utan problem.

5.4 Utvecklingsbeskrivning VCC

Inom vision 2020 vet vi att mer än 90% av olyckor kan härledas från förarfel. Distraction är en bidragande orsak till dessa olyckor så det är av yttersta vikt att kontinuerligt sträva till att förbättra interaktionen med nya funktioner i bilen. Vi behöver förstå hur vi skapar kundvärde med hjälp av ny funktionalitet som uppkopplade funktioner för med sig och

bilens egna infotainmentsystem och tertiär interaktion med teknologier som tagits in i kupén av föraren eller andra passagerare (exempelvis smartphones eller tablets).

samtidigt skräddarsyr upplevelse för att minska distraktion från körning med dessa nya system.

Under projektet har VCC steg för steg byggt kunskap om hur kundvärde kan skapas genom användande av prediktiva funktioner i talsystem och hur distraktion kan minskas för återkommande use-case och förståelse av hur man kan realisera funktioner i bilmiljö.

Tre övergripande faser har styrts arbetet:

1. Analys av kundbeteende
 - a. Data-mining av kundbeteenden från NHTS 2009 för kvantitativa analys och klustring av körmönster och definiera relevanta in-data för event-logging
2. Innovation av use case och provning av funktioner
 - a. Workshops med fokus att definiera vilka funktioner som skall implementeras
 - b. Utvärdering av prestanda av implementerade funktioner i Talkamatic platform
3. Utveckling av prototyp platform och anpassning till implementation i bil
 - a. Definition av lösningsarkitektur

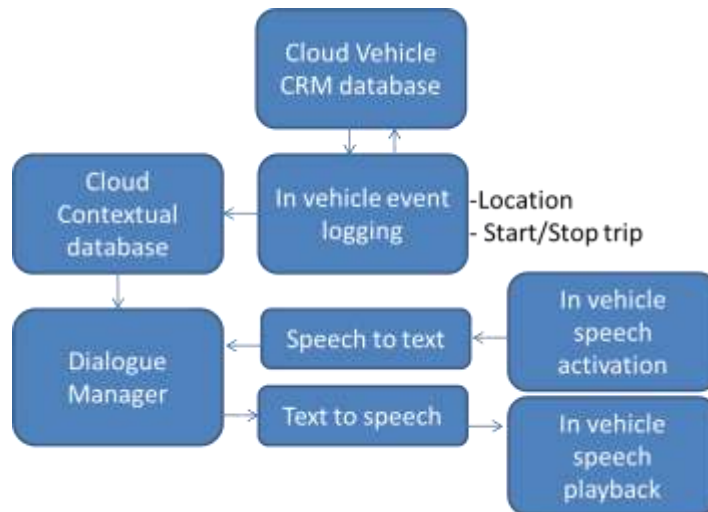


Figure 1 Lösningens arkitektur

- b. Utveckling av IMX-6 prototyp platform för att fungera tillsammans med Talkamatic lösning
- c. Definition av fortsatt applikation av gemensamt framtagen prototyp i övriga utvecklingsprojekt på VCC

5.5 Bidrag till FFI-mål

5.5.1 **Implementera ett lärande system som reducerar mängden repetitiva och informationsfattiga yttranden**

Det huvudsakliga målet med detta projektet är att designa och implementera stöd för lärande i TDM. Detta mål adresserar både målet om en kompetitiv svensk fordonsindustri och målet med ett minskat antal dödade och skadade i trafiken.

5.5.2 **Bygga UX-kunskap kring prediktiv dialog**

Detta mål adresserar både målet om en konkurrenskraftig svensk fordonsindustri och målet med ett minskat antal dödade och skadade i trafiken.

Projektet har bidragit till i huvudsak två kunskapsutvecklande block inom ramen för interaktionsdesign och user experience (UX). Arbetet har mynnat ut i teoriutveckling om designprinciper för adaptiva dialogsystem i förarmiljö (manuskript för publicering på en vetenskaplig arena är under bearbetning). Dessutom har projektet bidragit till ökad praktisk kunskap om rapid prototyping för att effektivt kunna studera UX-kvaliteter på slutanvändare genom en specifik Wizard-of-Oz-metod och –miljö (manuskript under bearbetning). Denna kunskap bidrar till att konkreta produktionsförbättringar görs hos deltagande företag. Genom samarbetet i arbetet med UX-kvaliteter mellan VCC, Talkamatic och Högskolan i Halmstad har projektet har dessutom bidragit till målet om stärkt samverkan mellan fordonsindustri och högskola.

5.5.3 **TDM på embeddedplattform**

Att kunna köra TDM på en embeddedplattform är en förutsättning för att kunna ta in Talkamatics teknik i bilen. För att projektresultatet skall kunna nyttiggöras inom den svenska fordonsindustrin fullt ut är detta en nödvändighet. Resultatet adresserar därmed målet om en kompetitiv fordonsindustri.

5.5.4 **Kunskapsbyggande om att ta in ny funktionalitet från mobilplattform in i bilen.**

Riktat sig mot målet om minskat antal dödade och skadade i trafiken.

5.5.5 **Patentansökan**

En preliminär svensk patentansökan är inlämnad av Talkamatic AB. Patentansökan avser den tekniska lösningen för att integrera kunskap med av olika styrka från en användarmodell i en dialog.

6 Spridning och publicering

6.1 Kunskaps- och resultatspridning

Talkamatic har kommit i kontakt med en grupp på Volvo Cars som arbetar med ruttprediktion. Det är rimligt att anta att arbetet som gjorts i detta projekt och i SSK kan berika varandra. Talkamatic arbetar också med Volvo Cars för att hitta en fortsättning på arbetet med prediktiv dialog. Det är också möjligt att resultaten från projekten integreras i prototypen som byggs i projektet SIMSI (FFI).

6.2 Publikationer

- Holdaj Pettersson, K., Thunberg, M. 2013. Adaptiva talbaserade system i fordon: designförslag för att främja user experience. Kandidatuppsats, Högskolan i Halmstad.
- Wärnestål, P., Kronlid, F. 2013. Towards a User Experience Design Framework for Adaptive Spoken Dialogue in Automotive Contexts. *Manuscript in preparation.*
- Kronlid, F., Wärnestål, P. 2013 Integrating Uncertain Information in an Information State Update Dialogue Manager. In preparation.

7 Slutsatser och fortsatt forskning

Vi har byggt en teknisk lösning som, vid upprepade användningsmönster, förenklar och förkortar dialogen, så att tiden som användaren är upptagen med dialog förkortas. Emellertid är det inte belagt att detta faktiskt reducerar distraktionsgraden hos användaren. Detta skulle vara intressant att undersöka i fortsatta projekt. I och med att en simulatormiljö med avancerade möjligheter att undersöka distraktion under utförande av sekundäruppgifter har byggts upp vid Centre for Language Technology/GU i samband med SIMSI-projektet, finns det goda möjligheter att utföra dessa studier.

Den tekniska lösningen som byggts i projektet modellerar en användares svar på systemets frågor, och inte användarens preferenser. Detta kan tyckas vara en yttlig skillnad, men kan vara betydelsefull i vissa fall, och kan leda till att systemet ställer frågor som det hade kunnat besvara själv om den hade modellerat användaren på ett smartare sätt. Det hade varit intressant att vidareutveckla användarmodeller så att användarens preferenser hade modellerats, för att se om man kan sänka användarens distraktioner ytterligare.

Ur ett designperspektiv är det tydligt att det saknas designprinciper och standarder för adaptiv dialoginteraktion i förarmiljö. Ur både säkerhets- och upplevelseperspektiv är ett sådant arbete viktigt. Detta arbete har, baserat på empiri från autentisk brukarsituation,



FORDONSSTRATEGISK
FORSKNING OCH INNOVATION

börjat skrapa på den ytan. Framtida forskning krävs för att detaljera ett mer komplett designramverk.

8 Deltagande parter och kontaktpersoner



Talkamatic AB, Fredrik Kronlid, fredrik@talkamatic.se, +46703602190



Volvo Cars, Jonas Söderqvist, jsoderqv@volvocars.com, +1 805 603 0997



Högskolan i Halmstad, Pontus Wärnestål, pontus.war nestal@hh.se, +4635167696