



Produkt- og Designpsykologi
Aalborg Universitet
<http://www.aau.dk>

AALBORG UNIVERSITET

STUDENTERRAPPORT

Titel:

Auditiv nudging ved talesystemer

Tema:

Kandidatspeciale

Projektperiode:

Forårssemestret 2016

Projektgruppe:

1080

Deltagere:

Anders Busch
Nima Nickolaj Mouritzen

Vejleder:

Lars Bo Larsen

Oplagstal: 4

Sidetal: 71

Afleveringsdato:

16. september 2016

Abstract:

The ability to converse with machines is something that has been long underway, and it is still not perfect, even though improvements are being made every year. However, there are still some ways in which it can improve, until it is made possible to utilise it the way it is intended. This study aims to investigate whether it is possible to nudge subjects, to use certain keywords in a conversational user interface, by using only auditory directed dialog and feedback, in different situations. This investigation was conducted by simulating a speech system in which subjects had to perform a number of different exercises in three different scenarios. The speech would use the keywords throughout the investigation, and the data hereof shows that using directed dialog nudges the subjects to use the keywords more than the use of feedback in a speech system. Furthermore, there is a significant difference in which words are being nudged the most, which could be the result of the context of the exercise in which they were used. Therefore when designing a speech system, one has to consider which words to use in a certain context, and to what degree they can be used in directed dialog, without it hampering the user experience.

FORORD

Dette kandidatspeciale er udarbejdet ved Det Teknisk-Naturvidenskabelige Fakultet på Aalborg Universitet og er skrevet i perioden fra den 15. juni til den 15. september 2016. Specialet er udarbejdet af gruppe 1080 bestående af to studerende på uddannelsen Produkt- og Designpsykologi.

Gruppen har haft Lars Bo Larsen som vejleder. I forbindelse med det udførte forsøg er der ydet hjælp fra bekendte af gruppen. Der skal lyde en stor tak til alle involverede parter.

INDHOLD

Indhold	4
1 Indledning	7
1.1 Tidlige talesystemer	7
1.2 Talesystemer i dag	8
2 Problemanalyse	9
2.1 Affordance	9
2.2 Tekst og tale	10
2.3 Nudging	11
3 Forsøgsbeskrivelse	13
3.1 Talesystemet	13
3.2 Forsøgsgennemgang	15
3.3 Forsøgsopstilling	16
3.4 Forsøgsrunderne	17
3.5 Skalaer	20
3.6 Exitinterview	20
3.7 Forsøgshypoteser	21
3.8 Pilotforsøg	22
4 Analyse	23
4.1 Resultater af nudging	24
4.2 Resultater af skalaerne	30
4.3 Besvarelser fra exitinterviewet	31
5 Diskussion	35
5.1 Forsøgspersonernes anvendelse af keywords	35
5.2 Forsøgsrundernes indflydelse	35
5.3 Forekomst af keywords	37
5.4 Skalaer på tværs af runder	37
5.5 Exitinterviewene	38
6 Konklusion	41
7 Perspektivering	43

7.1	Videre design af talesystemer	43
7.2	Videreundersøgelse	44
Litteratur		45
A	Forsøgsdokumenter	47
A.1	Materialeliste	47
A.2	Forsøgsprogram	48
A.3	Introduktion	50
A.4	Samtykkeerklæring	52
A.5	Exitinterview	53
A.6	Opgaverne til scenarie A - Træningsrunden	54
A.7	Opgaverne til scenarie B	56
A.8	Opgaverne til scenarie C	58
A.9	Talespor til scenarie A - Træningsrunden	60
A.10	Talespor til scenarie B	63
A.11	Talespor til scenarie C	68

INDLEDNING

Igennem de sidste mange årtier har mennesket drømt om at kunne kommunikere med maskiner på samme måde, som to personer kommunikerer med hinanden netop igennem *taleinteraktion*. Mennesket har nemlig siden en tidlig alder, anvendt sproget til at udtrykke dens ønsker og tilstand, og det har derfor også været drømmen, at kunne gøre have den selv samme form for kommunikation med en maskine, da det forekommer så naturligt (McTear, 2004). Det er derfor ikke så underligt, at nogle forfattere og filmskabere har igennem litteratur og film indenfor genren science-fiction kommet med deres egne bud på, hvordan denne interaktion mellem menneske og maskine potentiel kunne foregå, som for eksempel i filmen Rumrejsen 2001. Her bliver seeren introduceret for HAL, som er en computer med kunstig intelligens, og som er i stand til at kommunikere med mennesker netop igennem tale, og sågar føre en samtale. Dog er dette kun science-fiction, men idéen er ikke lang fra en realitet idag.

1.1 Tidlige talesystemer

En ting er at lave et system som er i stand til at tale til en person, og en anden ting er at lave et system som kan modtage taleinput fra en person. I 1950'erne og 1960'erne kom nogle af de første systemer, som var i stand til at modtage input via stemmestyring. Disse var dog så simple, at de kun kunne forstå cifre og nogle enkelte engelske ord, hvor det mest populære system blev kaldet "Audrey" (Pinola, 2011). Det tog dog for alvor fart i 1970'erne, hvor mere forskning var med til at gøre det muligt at udvikle systemet "Harpy" (Pinola, 2011). Harpy var i stand til at forstå 1011 ord, hvilket ikke lyder af meget med idags standarder, men dengang var det, der var muligt med den computerkraft som var til rådighed dengang (McTear, 2004). Det var derfor også meget begrænset dengang, hvad disse talesystemer blev anvendt til, men i løbet af 1980'erne begyndte de så at blive anvendt til kommercielt brug (Pinola, 2011) Dette inkluderede blandt andet data input til formular udfyldelse, eller at være i stand til at kontrollere apparater

(McTear, 2004). Alle disse eksempler led dog under de begrænsninger der var den teknologiske udvikling, hvilket betød at de var begrænset af de mængde ord en computer kunne processere og lagre. Derfor var talesystemerne ofte dedikerede systemer, som kun var i stand til at udføre nogle meget specifikke opgaver.

1.2 Talesystemer i dag

Idag ser talesystemer dog helt anderledes ud. I løbet af de sidste fire år er talesystemerne gået fra at have en meget simple og direkte brug, til en mere avanceret funktionalitet, nemlig som *conversational agents* (vil fremover blive refereret til som CA), som kan tilgås fra din lomme eller i dit hjem. Systemer ligesom “Siri”, “Google Now” og “Cortana” er alle CA'er som kan tilgås via en smartphone, og de kan meget mere end hvad de gamle dedikerede talesystemer, blandt andet fordi teknologien er til det. Luger og Sellen (2016) mener at denne form for interaktion har potentiale til at forbedre brugeroplevelsen indenfor mennesker-computer interaktion (HCI), da taleinteraktionen åbner op for en mere direkte måde at interagere og tilgå funktioner på. På grund af den teknologiske udvikling er det også blevet muligt at opnå en mere menneskelig interaktion med et talesystem. Faktisk menes det at under en social interaktion responderer mennesker på samme måde for menneske-computer interaktion, som når det er menneske-menneske interaktion (Reeves og Nass, 1998). Dog kan en alt for menneskelig feedback fra et talesystem sætte nogle urealistiske forventninger til hvad talesystemet egentlig er i stand til, da nogle personer begynder at opfatte for eksempel en CA, som havende samme menneskelige færdigheder (McTear, 2004).

I en undersøgelse udført af (Luger og Sellen, 2016), hvor brugen af en håndfuld CA'er bliver undersøgt af personer som har erfaring med talesystemer, finder de frem til at der stadig er en række problemer med interaktionen med talesystemer. Respondenterne i denne undersøgelse beretter om, at når de initierende giver talesystemet en kommando, så anvender de deres naturlige sprog, men hvis eller når dette ikke virker, simplificerer de det indtil at systemet forstår hvad der bliver sagt. Dette har så ført til at nogle af respondenter begynder at fravælge at bruge andre ord end de keywords som talesystemet respondere til. Det kunne derfor være interessant at se på, om det er muligt for talesystemet at gøre disse keywords mere tydelige for brugeren, og påvirke dem til at bruge dem. Dette kan derfor lede op til en initierende problemformulering som lyder:

Hvorvidt er det at påvirke brugere til at anvende specifikke ord til at interagere med et talesystem?

PROBLEMANALYSE

2.1 Affordance

Affordance blev først introduceret af James J. Gibson i 1979 og han beskrev affordance således:

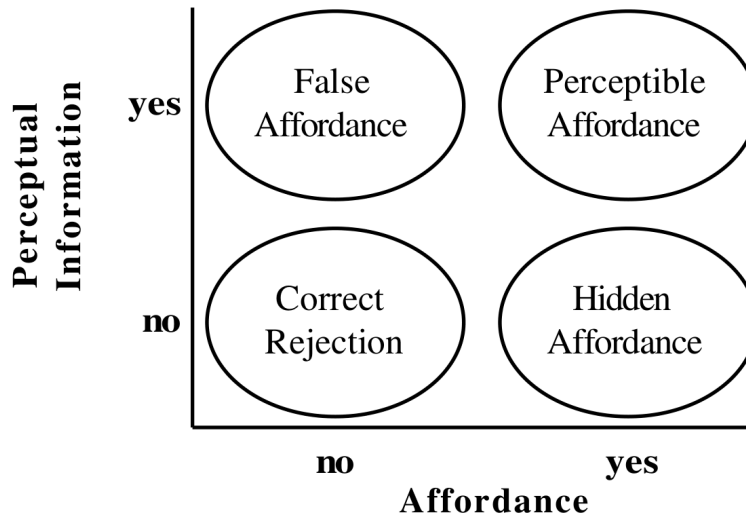
The affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill. ... I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment. (Gibson, 1986)

Sidenhen er udtrykket affordance blevet et buzzword indenfor HCI, især på grund af Donald Normans overførsel af affordance til GUIs. Han har dog nu ændret sin beskrivelse af affordance, og sidst har han valgt at droppe ordet affordance og i stedet kalde det signifiers.

William Gaver har beskrevet forholdet mellem hvorvidt en affordance eksisterer ved et objekt, og om der er information om affordancen, som kan opfattes, hvilket kan ses i hans model på Figur 2.1.

Ud fra modellen kan der ske fire forskellige udfald i en given situation. Hvis et objekt har en affordance og informationen omkring denne affordance er perciperbar, så vil det være en affordance, der fungerer. Hvis objektet har en affordance men informationen derimod ikke er perciperbar, så vil det være en skjult affordance, da affordancen ikke kan opfattes. I det omvendte tilfælde, hvor der er information om en affordance, men denne affordance ikke er til stede ved objektet, så er det en falsk affordance. I det sidste tilfælde er der hverken affordance eller information om denne til stede ved objektet, og dette kalder Gaver korrekt afvisning.

Affordance er altså de muligheder, mennesker opfatter (og muligvis overser), og disse kan bruges til at afgøre, hvilke måder et problem kan klares på.



Figur 2.1: Gavers model over sammenhængen mellem affordance og tilgængelig information omkring affordancen (Gaver, 1991).

2.2 Tekst og tale

I 1991 udgav Elizabeth Zoltan-Ford en artikel omhandlende et forsøg, hvor hun fik en række forsøgspersoner til at skulle interagere med en computerterminal. Hun begrænsede nogen brugere til at skulle svare på samme måde som computeren spurgte på og andre lod hun svare, som de ville. Samtidig lod hun en gruppe skrive deres svar ind via et tastatur og en anden gruppe lod hun svare igennem et talesystem. Selvom hun snød sine forsøgspersoner til at tro, at systemet rent faktisk virkede, hvor der i virkeligheden sad en forsøgsvejleder og skrev svarene tilbage til forsøgspersonerne, så fandt hun frem til forskellige ting i hendes forsøg (?).

Hun fandt ikke overraskende ud af, at det var mere effektivt at få forsøgspersonerne til at ændre deres input ved at bruge fejlmeddelelser og gentagelser frem for, at forsøgspersonerne selv formede deres input til computeren efter, hvad den spurgte efter. Hun fandt dog også ud af, at der ikke var forskel fra påvirkningen af computeren på, hvorvidt forsøgspersonerne skulle angive deres svar via tastaturet eller ved hjælp af stemmekommandoer. Det tyder derfor på, at forsøgspersoner ændrer deres sprogbrug både i skrift- og talesprog, selv når de kun bliver påvirket visuelt. Hun fandt også ud af, at de forsøgspersoner, som skulle give input via tale tog opgaverne som en samtale, og lod computeren spørge efter informationen, hvor dem, der skulle give input via tastaturet forsøgte at få al information ned i én sætning (?).

Hun fandt også ud af, at de personer, som kunne svare på deres egen måde, følte sig mere tilfreds med computere end dem, som var begrænset til at svare på en bestemt måde. (?).

Hun foreslår tre ting, der bør tages højde for, når sådanne systemer skal designes:

- Giv konsistent output, brugere vil forme sig efter det
- Design programmet til at kommunikere kort, brugere vil forme sig mere efter korte kommandoer
- Hav neutrale fejlbeskeder, der gentager det relevante sprogbrug

Ifølge ?, så ændrer mennesker deres sprogbrug efter, hvilken situation de er i. Mennesker har en række variationer af deres sprogbrug med forskellige rytmer og vaner, som bliver bragt frem og lagt væk alt efter, hvornår det virker idéelt for den enkelte person (?).

2.3 Nudging

Nudging beskrives som en måde at påvirke på, hvori mulighederne ikke begrænses, men derimod forsøger det at "guide" hen mod en bestemt mulighed (?). Ifølge (?) er nudging en måde at undgå de dårlige valg, da denne artikel fokuserer på sundhed. Fælles er, at nudging er en måde at påvirke på i en bestemt retning, som umiddelbart ikke er uønsket for den påvirkede. Eksempler på nudging i gadebilledet er fodspor malet på asfalten, der leder hen mod skraldespande. Idéen er her, at se disse fodspor skulle give en hentydning til, at skrald kan smides ud i skraldespanden for enden af sporet, og sporene giver indikationen af, at en anden har gjort det før dig.

Affordance og nudging ligger ret tæt på hinanden, emnemæssigt, dog er de helt forskellige. Affordance handler om hvilke muligheder, der er tilgængelige, hvor nudging prøver at vise, at en af mulighederne er bedre at vælge end de andre. Nudging fungerer som en indirekte påvirkning, da en mulighed ikke må virke bedre end andre ud fra, hvor billig, tidsbesparende eller hvor meget nemmere den er end de andre muligheder, den skal blot virke som det bedre valg blandt ligeværdige muligheder.

Det vælges til det kommende forsøg at benytte nudging til at påvirke forsøgspersonerne. Derfor opstilles følgende problemformulering:

Hvorvidt er det muligt at nudge brugere auditivt til at anvende specifikke ord til interaktion med et talesystem?

Til at nudge med vælges det at benytte to emner, der normalt bruges inden for talesystemer, initiativstyring og bekræftelse.

Initiativstyring er klassificeringen for de former for styring, der normalt benyttes ved talesystemer. Den udgave af initiativstyring, der vælges til forsøget kaldes *Directed Dialog*, hvilket betyder, at talesystemet styrer interaktionen ved at den stiller et spørgsmål, og forsøgspersonen skal blot svare på spørgsmålet (Hempel, 2010). Hvis forsøgspersonen skulle give mere information end denne blev spurgt efter, da vil systemet enten blot tage imod den efterspurgte information, eller systemet vil spørge efter informationen igen. Det vælges dog at kalde denne form for styring for *initiativstyring* fremover. Grunden til, at initiativstyring er valgt, er fordi, at det kan sikres, at alle forsøgspersoner kommer igennem samme forløb, hvis systemet styrer interaktionen. Derudover vil der blive brugt specifikke keywords i initiativstyringen, som forhåbentligt vil nudge forsøgspersonerne til at benytte dem.

Bekræftelse vil blive brugt til at forsøge at nudge forsøgspersonerne efter, at de har udført en opgave korrekt. Dette vil blive gjort ved at talesystemet bekræfter svaret og gentager det specifikke keyword, som er valgt til den enkelte opgave.

FORSØGSBESKRIVELSE

Til undersøgelsen af problemformuleringen opstilles et forsøg hvor forsøgspersoner vil gennemgå nogle forskellige opgaver med henblik på anvendelsen af et talesystem. Formålet med forsøget vil være at prøve at nudge forsøgspersonerne til at anvende specifikke ord til de respektive funktioner der anvendes i opgaverne, og få feedback på anvendelsen af oplevelsen med talesystemet. Dette leder op til en overordnet hypotese som lyder:

Forsøgspersoner kan nudges til at tilpasse deres sprogbrug til det ordvalg et talesystem anvender.

Forsøget vil komme til at bestå af to runder og en træningsrunde, som hver især vil indeholde tre opgaver, og forsøget vil foregå med et within-subject design. For at undersøge hypotesen vil forsøgspersonens forsøgsgang blive videooptaget igennem hele forsøget. Deres ordvalg igennem de tre opgaver for hver runde vil blive sammenlignet med de prædefinerede ord, som talesystemet anvender. Dette gøres for at se på om forsøgspersonerne vælger at anvende samme ordvalg som talesystemet, eller om de ændrer deres ordvalg fra en opgave til den anden. Hvilke ord der vil blive lagt vægt på i forsøget er defineret i afsnit 3.4.

3.1 Talesystemet

I forsøget bliver der ikke brugt et selvfungerende talesystem, som kan forstå og selv svare på de inputs, som forsøgspersonen måtte give. Dette vil dog fremover blive referet til som “talesystemet”. Talesystemet vil i stedet blive simuleret ved hjælp af en tale-til-tekst-hjemmeside, *Ivona*¹. Svarene er lavet på forhånd, og bliver afspillet gennem en højtaler, som vil stå foran forsøgspersonerne. Der vil ikke være nogen form for GUI foran forsøgspersonen, da det ønskes at nudge

¹<https://www.ivona.com/>

forsøgspersoner udelukkende ved hjælp af talesystemets audio, som var det en CA. For at undgå at forsøgspersonerne tænker, at det er en computer, der afspiller talesystemets kommandoer, vil højttaleren være maskeret, så den ikke ligner en højttaler. Maskering består af nogle A4 ark med et print af en Amazon Echo², for at give idéen omkring et rigtigt talesystem (se Figur 3.1). Forsøgslederen, som sidder og vælger, hvilke kommandoer, talesystemet skal afspille, vil sidde i et kontrolrum, så forsøgspersonerne ikke ved, at det er en person, som styrer talesystemet.



Figur 3.1: Figuren viser højttaleren til talesystemet, med og uden maskering.

Talestruktur

For at kunne simulere et talesystem i forskellige situationer, skal der defineres en talestruktur for talesystemet. Til talestrukturen vælges det at anvende variationer mellem et initiativstyret talesystem, et talesystem som kun giver feedback til brugeren ved at bekræfte en handling ved gentagelse, og et talesystem som anvender begge dele. Ved at variere de forskellige former for talestruktur er det muligt at se på om der er en forskel på hvorvidt forsøgspersoner bliver nudget ud fra disse talestrukturer.

Initiativstyring anvendes for at præsentere funktionerne for brugeren samtidig med at forsøge at nudge dem til specifikke keywords til funktionerne. Dette anvendes i scenarie B og C da dette er meget brugt i slotfiller situationer, for at sikre, at brugeren gennemgår de nødvendige trin for at opnå et specifikt mål. Da initiativstyring er godt til nye brugere, da det sikrer, at al nødvendig information bliver indtastet i systemet, og brugeren får fornemmelsen af, at systemet “holder dem i hånden” under hele interaktionen, kan det derimod være til gene for erfarne brugere da de kan se dette som et irritationsmoment, da de ved, hvilket information er nødvendigt, og ikke har behovet for hjælp som nye brugere.

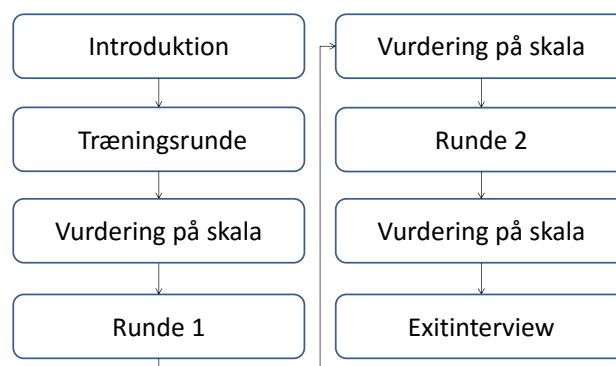
Bekræftelse anvendes for at give brugeren feedback på, hvad systemet har modtaget af input, samtidig med at det også forsøger at nudge forsøgspersonerne til at benytte specifikke input til funktionerne. Det fungerer derfor som et slags facit for brugeren, og det kan derfor antages at brugeren vil være modtagelig overfor nudging igennem denne form for talestruktur, da dette kan forekomme som det rigtige “resultat”.

²Billedet anvendt som maskering stammer fra https://echosim.io/static/assets/img/amazon_echo.png

Til selve forsøget vil talesystemet godkende alle de former for variationer af input en forsøgsperson måtte komme med, så længe disse giver mening i forhold til opgaven. Dette er dog noget som forsøgslederen skal vurdere mens han styrer systemet. Forsøgslederen vil derfor sidde med en liste over talespor for de forskellige scenarier, som vil hjælpe i vurderingen om hvorvidt et input vil blive godkendt af talesystemet. Udover talesystemets talespor til hver enkelt scenarie vil være et generelt lydclip som forsøgslederen kan drage nytte af under forsøget. Dette er “Jeg forstår ikke din kommando. Prøv igen”, hvilket vil blive anvendt når en forsøgspersons input ikke er godkendt.

3.2 Forsøgsgennemgang

Forsøget er udformet som et within-subject design, hvor forsøgspersonerne vil blive præsenteret for en træningsrunde og to forskellige scenarier med fokus på anvendelsen af et talesystem. Da talesystemer anvendes i mange forskellige sammenhænge (slot-fillers, CA), vælges det at forme scenarierne forskelligt, med forskellige former for styring. Dette gøres for at undersøge om de forskellige former for styring har en effekt på hvorvidt om brugeren bliver nudget af talesystemet, samt om det har en effekt på brugerens oplevelse af talesystemet. Derudover vælges det at have tre lignende opgaver til hvert scenarie, for således at se om nudgingeffekten overføres fra opgave til opgave. Til hvert scenarie vil der være nogle udvalgte udtryk (*keywords*), der vil blive lagt vægt på i forhold til om forsøgspersonen tilpasser sit sprogbrug til det, som talesystemet anvender. Disse vil blive defineret under beskrivelsen af de respektive scenarier. En illustration af forsøgsprocessen kan findes på Figur 3.2.



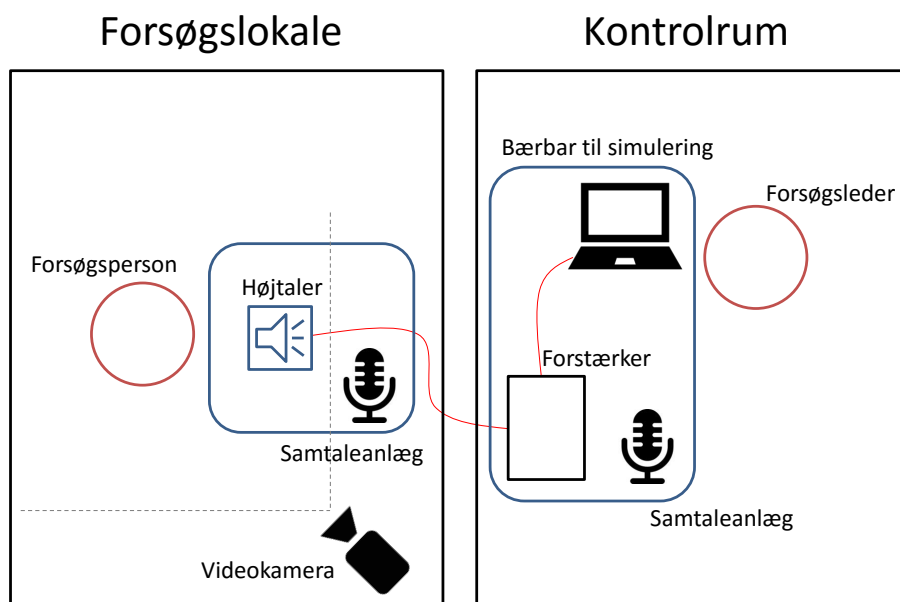
Figur 3.2: Flowchart over forsøgsforløbet.

Forsøgspersonen vil i starten af forsøget få udleveret et opgavesæt, hvor der vil være en introduktion til forsøget, samt beskrivelser af alle scenarierne og de dertilhørende opgaver. Når forsøgspersonen har læst introduktionen vil de blive bedt om at underskrive en samtykkeerklæring, for at sikre at dataet må anvendes til projektet. Efterfølgende skal forsøgspersonerne gennemføre en træningsrunde for at gøre dem fortrolig med opgaven. Første runde begynder

først når forsøgspersonen er klar til at fortsætte. Scenarierne vil komme i randomiseret rækkefølge for hver forsøgsperson, for at mindske risikoen for en rækkefølgeeffekt. Når første runde er afsluttet, vil forsøgspersonerne blive bedt om at vurdere deres oplevelse på nogle skalaer, som findes i opgavesættet. Når forsøgspersonen er færdig med alle runderne, vil forsøget blive afsluttet med et exitinterview.

3.3 Forsøgsopstilling

I dette afsnit vil selve forsøgsopstilling kort blive beskrevet, samt hvilke overvejelser der ligger bag. Til forsøget vil forsøgspersonen sidde alene ved et bord i et stille og roligt rum. På bordet vil talesystemet, en maskeret højttaler, samt et samtaleanlæg stå. Samtaleanlægget anvendes for at forsøgslederen kan høre hvad der bliver sagt i rummet, så han kan styre talesystemet ud fra det forsøgspersonen siger. Samtaleanlægget sender kun lyd den ene vej, så forsøgspersonen kan ikke høre forsøgslederen, medmindre forsøgslederen ønsker det. Forsøgslederen vil sidde i et kontrolrum ved siden af, hvor han styrer det, som talesystemet skal sige ved hjælp af en bærbar, som indeholder alle de nødvendige lydfiler. Et eksempel på forsøgsprogrammet kan ses på Bilag A.2. Bærbaren sender lyden ud til højttaleren via en forstærker, som står i kontrolrummet. Hele forsøget vil blive filmet, da det er ud fra disse optagelser at dataet vil komme af. Hele forsøgsopstillingen kan ses på Figur 3.3.



Figur 3.3: Illustration af forsøgsopstillingen.

3.4 Forsøgsrunderne

I forsøget vil der være tre forskellige scenarier, som forsøgspersonerne skal gennemgå, hvoraf den ene er en træningsrunde. For at tilpasse scenarierne til den styreform der ønskes at undersøge, vil de anvendte scenarier simulere et talesystem i forskellige sammenhænge. Dette er for at opsætte nogle forholdsvis reelle scenarier, hvori at en bestemt form for talestruktur måtte forekomme, da et talesystem kan fungere anderledes i et slot-fillersystem end det gør ved en CA. Derudover er det også muligt at få forsøgspersonens vurdering af talesystemet i forskellige situationer. Til hvert scenarie vil der være tre overordnet opgaver, for at se om effekten af nudging fremkommer fra den ene opgave til den anden. Hver opgave vil blive præsenteret på et opgaveark, som forsøgspersonen vil få udleveret i starten af forsøget. Her vil opgaven blive beskrevet med så lidt information som muligt, dog nok til at forsøgspersonen kan udføre opgaven, for at undgå at præge dem allerede her, da idéen er, at talesystemet i sig selv skal nudge forsøgspersonen i en bestemt retning. Forsøgslederen vil sidde med en liste af forskellige input for at vurdere, hvorvidt om forsøgspersonen har givet et godkendt input, med eller uden brug af keywrodet, hvoraf han tillader forsøgspersonen til at gå videre til næste skridt i opgaven. Hvert scenarie vil have sit eget sæt af keywords, da scenarierne er for forskellige til at anvende de samme sæt af keywords. På denne måde er det også muligt at teste, hvorvidt om ordet har en effekt på, om en forsøgsperson bliver nudget eller ej.

Scenarie A - Træningsrunden

Træningsrunden omhandler en form for intelligent hjemmekontrol. Forsøgspersonerne vil blive bedt om at skulle ændre på indstillinger i hjemmet, såsom at rulle gardiner for/fra, ændre temperaturen i rummet via en termostat, ændre belysningen i rummet og ændre baggrundsmusikken. De tre opgaver lyder i dette scenarie:

1. Opgave: Tænd lyset i køkkenet, indstil temperaturen til 24 grader, sæt rockmusik på og træk gardinerne fra i soveværelset.
2. Opgave: Tænd lyset i stuen, indstil temperaturen til 19 grader, sæt hiphopmusik på og træk gardinerne fra i køkkenet.
3. Opgave: Sluk lyset i soveværelset, indstil temperaturen til 22 grader, sæt jazzmusik på og træk gardinerne for i stuen.

Beskrivelserne, som forsøgspersonerne bliver præsenteret for i forsøget, kan ses i Bilag A.6. Talesystemets talespor kommer til at indeholde keywords, som kan anvendes indenfor de relevante funktioner. De keywords, der vælges at fokusere på i dette scenarie, er:

- Lamper
- Indstil
- Afspil

- Ruller for/fra

Det vælges at fokusere på ord som **Indstil** og **Afspil** som keywords, da det antages at disse ord ofte anvendes i det scenarie de præsenteres i, hvor der også findes mange andre måder at formulere samme betydning. Ordet **Ruller** er inkluderet som keyword i sammenhængen med gardiner, da det antages at de fleste personer normalt ville anvende “Trækker”, da dette er en mere direkte beskrivelse af den handling der laves, og det kunne derfor være interessant at se, om det er muligt at nudge forsøgspersoner til at anvende det ord de måske normalt ikke bruger. Af samme grund vælges det at fokusere på ordet **Lamper** istedet for “Lys”.

Formålet med træningsrunden er at det skal være simpelt for forsøgspersonerne at forstå, derfor er det normale funktioner, som kan udføres i de fleste hjem. Der forsøges dog stadig at nudge forsøgspersonerne, dog uden nogen form for initiativstyring, og kun ved brug af bekræftelse. Dette er for at skabe et mere troværdigt scenarie i forhold til et intelligent hjem, da det ville virke underligt, hvis et intelligent hjemmekontrol ville foreslå at udføre en funktion, medmindre den blev spurgt om tilgængelige funktioner. Det vælges at anvende dette scenarie som træningsrunde, da det anses for at være det scenarie der stresser forsøgspersonerne mindst, ved at de selv er i fuld kontrol over hvad de kan/skal sige, og hvornår de siger det, uden at systemet tager initiativet fra forsøgspersonen. Hele talesporet til talesystemet, samt nogle af de godkendte svarmuligheder for forsøgspersonerne, kan ses i Bilag A.9.

Scenarie B

Scenarie B er opstillet med fokus på bestillingen af en flyrejse. Forsøgspersonerne vil blive bedt om at angive afgang, destination, tidspunkt, dato og det antal kufferter de skal have med på turen. De tre opgaver forsøgspersonerne skal lave til dette scenarie er:

1. Opgave: Bestille en rejse fra Amsterdam til New York, med afgang den 12. September 2016 klokken 06:45 med tre styks bagage.
2. Opgave: Bestille en rejse fra Frankfurt til Shanghai, med afgang den 25. September 2016 klokken 12:30 med et styks bagage.
3. Opgave: Bestille en rejse fra Aalborg til London, med afgang den 2. Oktober 2016 klokken 17:00 med to styks bagage.

Beskrivelserne som forsøgspersonen bliver præsenteret for i forsøget kan ses i Bilag A.7. Talesystemets talespor kommer til at indholde relevante ord som normalt anvendes indenfor emnet “rejser”. De keywords der vælges at fokusere på i dette scenarie er:

- Destination
- Afgang
- Afrejse
- Månedens navn frem for tal

- Styk bagage

Det vælges at fokusere på ord som **Destination** og **Afgang** som keywords, da det antages at disse ord ofte anvendes i det scenarie de præsenteres i, hvor der også findes mange andre måder at formulere samme betydning, dog uden at bruge dagligdags ord som "Til og Fra". Ordet **Afrejse** er inkluderet som keyword i sammenhængen med tidspunkt for en rejse, da det antages at de fleste personer normalt ville anvende et udtryk som "Tage afsted", da det antages at dette er et mere alment udtryk, og det kunne derfor være interessant at se, om det er muligt at nudge forsøgspersoner til at anvende det ord de måske normalt ikke bruger. Af samme grund vælges det at fokusere på udtrykket **Styk bagage** istedet for "Tasker" og "Kufferter". Derudover vælges det at se om det er muligt at nudge forsøgspersonerne til at sige månedens navn, selvom den i opgaven vil stå beskrevet med tal. Dette er mest for at observere om de vælger at anvende den nemme løsning, og bare læse op fra opgaven, eller om de tilpasser det talesystemet.

Målet er som nævnt at nudge forsøgspersoner til at anvende disse keywords igennem dette scenarie, dog er det ikke de eneste udtryk, som vil blive godkendt for at forsøgspersonen kan komme videre i opgaven. Talesystemet vil guide forsøgspersonen hele vejen igennem processen af bestillingen af en rejse, hvor det tager styringen af hvad personen skal sige, og be- eller afkræfter det efterfølgende. Hele talesporet til scenarie B, samt nogle af de godkendte svarmuligheder for forsøgspersonerne, kan ses i Bilag A.10.

Scenarie C

Scenarie C er opsat med fokus på oprettelsen af begivenheder i en digital kalender. Her skal forsøgspersonerne angive indholdet af begivenheden, dato, tidspunkt og en påmindelse tilhørende begivenheden. De tre opgaver forsøgspersonerne skal udføre til dette scenarie er:

1. Opgave: Lav en begivenhed til den 12-10-2016, klokken 18:00, hvor emnet er en 20-års fødselsdag, og påmindelsen er: "købe gave".
2. Opgave: Lav en begivenhed til den 07-12-2016, klokken 13:45, hvor emnet er et bryllup, og påmindelsen er: "stryge tøj".
3. Opgave: Lav en begivenhed til den 27-05-2017, klokken 10:00, hvor emnet er karneval, og påmindelsen er: "finde kostume".

Beskrivelserne som forsøgspersonerne bliver præsenteret for i forsøget kan ses i Bilag A.8. Talesystemets talespor kommer til at indholde relevante ord som normalt anvendes indenfor begivenheder i en kalender. De keywords der vælges at fokusere på i dette scenarie er:

- Foregår
- Finder sted
- Emne
- Påminde

Det vælges at fokusere på ord som **Foregår**, **Påminde** og **Emne** som keywords, da det antages at disse ord ofte anvendes i det scenarie de præsenteres i, hvor der også findes mange andre måder at formulere samme betydning. Udtrykket **Finder sted** er inkluderet som keyword i sammenhængen med tidspunktet for en begivenhed, da det antages at de fleste personer normalt ville anvende “Klokkeslæt” eller “Tidspunkt”, da disse er en mere direkte beskrivelse af tid, og det kunne derfor være interessant at se, om det er muligt at nudge forsøgspersoner til at anvende det udtryk de måske normalt ikke bruger.

Ligesom i scenarie B er målet at nudge forsøgspersonerne til at bruge disse ord, her benyttes der dog ikke nogen gentagende bekræftelse af forsøgspersonernes svar. Hele talesporet til scenarie C, samt nogle af de godkendte svarmuligheder for forsøgspersonerne, kan ses i Bilag A.11.

3.5 Skalaer

Til forsøget er der opstillet fire skalaer for at få et bedre indblik i forsøgspersonernes oplevelse med talesystemet, og de opgaver de skal udføre til de forskellige scenarier. Alle skalaerne er opstillet som ti-punktsskalaer, med to polariserende ekstremer i hver ende af skalaen. Det blev vurderet at ti-punktsskalaer er passende til dette formål, selvom en VAS-skala giver forsøgspersonerne mulighed til at give en mere nuanceret vurdering. De fire skalaer anvendt til forsøget kan ses på Figur 3.4.

Disse skalaer er indsat i forsøget, for at undersøge om de forskellige scenarier er sammenlignelige, da for eksempel en variabel som sværhedsgrad kan have indflydelse på, hvordan forsøgspersonerne griber opgaven an. Disse skalaer vil forsøgspersonerne få udleveret på papir før forsøget, og de vil mellem hver runde blive mindet om, at de skal huske at give deres vurdering, før de kan gå videre til næste runde.

3.6 Exitinterview

Efter forsøget skal hver forsøgsperson igennem et exitinterview. Formålet med exitinterviewet er at finde ud af, om forsøgspersonerne bemærkede, at de blev nudget, og om det gjorde dem noget. Ligeledes udføres det for at høre, hvad de synes om forsøget, da deres oplevelse med forsøget kan have påvirket deres interaktion. De vil blive spurgt ind til følgende områder; hvordan de synes forsøget gik, hvad de synes om talesystemet og stemmen, som systemet brugte, om de har erfaring med talesystemer fra tidligere og om de valgte at ændre deres sprogbrug i forsøget. Forsøgspersonerne vil blive bedt om at uddybe deres svar, hvis de nævner ting, som kunne være interessante for forsøget. Der vil derudover blive noteret demografiske oplysninger om forsøgspersonerne, alder, køn og studieretning, da dette muligvis kan være med til at forklare deres interaktion med talesystemet eller deres svar. De præcise spørgsmål, som bliver stillet i exitinterviewet, kan ses i Bilag A.5.

Hvor svært synes du det var at forstå opgaverne?

1 2 3 4 5 6 7 8 9 10

Meget svært Meget nemt

Hvordan var din oplevelse med at interagere med talesystemet?

1 2 3 4 5 6 7 8 9 10

Meget dårlig Meget god

Hvor pålidelig virkede talesystemet til at være?

1 2 3 4 5 6 7 8 9 10

Meget upålidelig Meget pålidelig

Hvor sikker var du i at anvende talesystemets funktioner?

1 2 3 4 5 6 7 8 9 10

Meget usikker Meget sikker

Figur 3.4: Denne figur illustrerer de skalaer, som forsøgspersonerne skal vurdere på efter hver runde.

3.7 Forsøgshypoteser

Til forsøget opsættes der to hypoteser, der ønskes at få svar på ud fra dataet. Disse hypoteser er opsat ud fra, at forsøget kommer til at besvare problemformuleringen og de to hypoteser lyder som følgende:

1. Det er nemmere at nudge ved at bruge både initiativstyring og bekræftelse end kun at bruge en af disse metoder
2. Nogle ord vil være nemmere at nudge end andre

Disse to hypoteser vil blive henvist til i databehandlingen, når det relevante data undersøges.

3.8 Pilotforsøg

Der blev udført to pilotforsøg, hvor der blev lavet ændringer efter hvert forsøg, for at forbedre setup'et til det endelige forsøg. Pilotforsøgene blev udført i samme lokale med samme fysiske opsætning som det endelige forsøg.

Første pilotforsøg

Dette pilotforsøg blev udført med en kvindelig studerende på 23 år fra Elektronik og IT på 5. semester. Pilotforsøget viste en række tekniske problemer, som der ikke var taget højde for.

Der blev fundet ud af, at ved kun at have én besked til når forsøgspersonerne lavede fejl havde konsekvenser. Hvis hun lavede en fejl ved at hun gav et for kort svar ville systemet give den samme fejlmeddelelse som hvis hun angav den forkerte destination. Dette gjorde at hun ændrede hendes sætninger markant fra et forsøg til det næste, hvilket ikke var hensigten, hvis hun var tæt på at få sagt sætningen korrekt. Derfor vælges der at lave en ekstra fejlmeddelelse, som lyder: "Jeg forstår næsten hvad du siger. Prøv med en hel sætning". Det antages, at denne sætning vil hjælpe forsøgspersonerne til at gøre deres sætning længere, men ikke nødvendigvis ændre på de allerede benyttede ord, hvorimod fejlmeddelelsen "Jeg forstår ikke din kommando. Prøv igen" skulle få forsøgspersonerne til at tænke over, hvad de havde sagt og få ændret på noget i sætningen. Der blev ud over den ekstra fejlmeddelelse også lavet en sætning fra talesystemet, som skulle minde forsøgspersonerne om, hvornår de skulle vurdere runderne på skalapapirerne. Denne blev afspillet i slutningen af hver runde.

Endvidere blev der i introduktionen til forsøget tilføjet et eksempel på, hvad der menes med en hel sætning, da pilotforsøgspersonen viste tvivl i forhold til, hvordan hun skulle sige tingene.

Andet pilotforsøg

Det andet pilotforsøg blev udført med en mandlig ph.d.-studerende på 25 år. Dette pilotforsøg viste en række forståelsesproblemer, som kunne fjernes ved at redigere i teksterne på de forskellige dokumenter, forsøgspersonerne blev givet inden starten på forsøget.

Der blev fjernet unødvendig information fra introduktionen og den blev samtidig gjort mere læsbar, samtykkeerklæringen fik tilføjet mere information om forsøgslederne og der blev tilføjet små sætninger på de forskellige papirer, som skulle hjælpe forsøgspersonerne til at guide dem videre til, hvad de skulle gøre.

Det blev vurderet, at det ikke var nødvendigt at have flere pilotforsøgspersoner med efter anden omgang, da der ikke var flere problemer til stede.

ANALYSE

Der deltog i alt 20 personer i forsøget, 13 mænd og 7 kvinder. Forsøgspersonerne var mellem 20 og 26 år og gennemsnitsalderen var på 22,5 år. Information om forsøgspersonerne kan ses i Tabel 4.1.

Forsøgsperson	Alder	Køn	Studieretning
1	22 år	Kvinde	PDP5
2	23 år	Kvinde	PDP5
3	23 år	Mand	PDP5
4	21 år	Mand	PDP5
5	22 år	Mand	PDP5
6	22 år	Mand	PDP5
7	22 år	Mand	PDP5
8	22 år	Mand	PDP5
9	21 år	Mand	PDP5
10	24 år	Kvinde	PDP5
11	22 år	Mand	PDP5
12	23 år	Kvinde	PDP7
13	24 år	Mand	PDP7
14	22 år	Kvinde	PDP7
15	20 år	Mand	PDP3
16	21 år	Mand	PDP3
17	23 år	Kvinde	PDP3
18	23 år	Kvinde	PDP3
19	25 år	Mand	PDP10
20	26 år	Mand	SW3

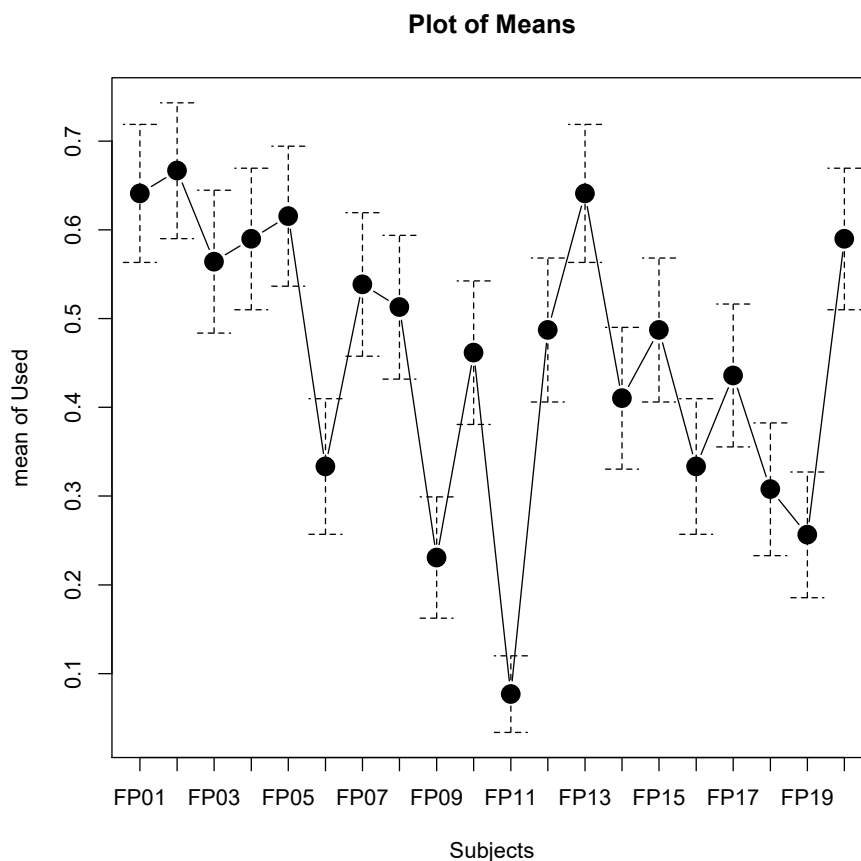
Tabel 4.1: Information om forsøgspersonerne. PDP står for Produkt- og Designpsykologi og SW står for Softwareingeniør. Tallet under studieretningen angiver det semester, som forsøgspersonen går på.

4.1 Resultater af nudging

Ved at se videooptagelserne igennem for hver forsøgsperson er der optalt, hvor mange gange hver person har brugt de forskellige keywords og hvornår i forsøget, de brugte dem. Der blev også optalt, hvor mange fejl hver forsøgsperson lavede. Alt data fra videoerne kan ses på bilags-cd'en i mappen "Forsøg" under navnet "Forsøgsdata".

Forsøgspersoners brug af keywords og fejl

Først ses der på forsøgspersonerne og hvordan de har klaret sig i forhold til hinanden. Dette gøres for at undersøge, om nogle forsøgspersoner skiller sig ud fra de andre. Ved at lave et Plot of Means over forsøgspersonerne og antallet af gange, hvor de har brugt de valgte keywords, kan der ses, om der er forskel forsøgspersonerne imellem. Dette plot kan ses på Figur 4.1.

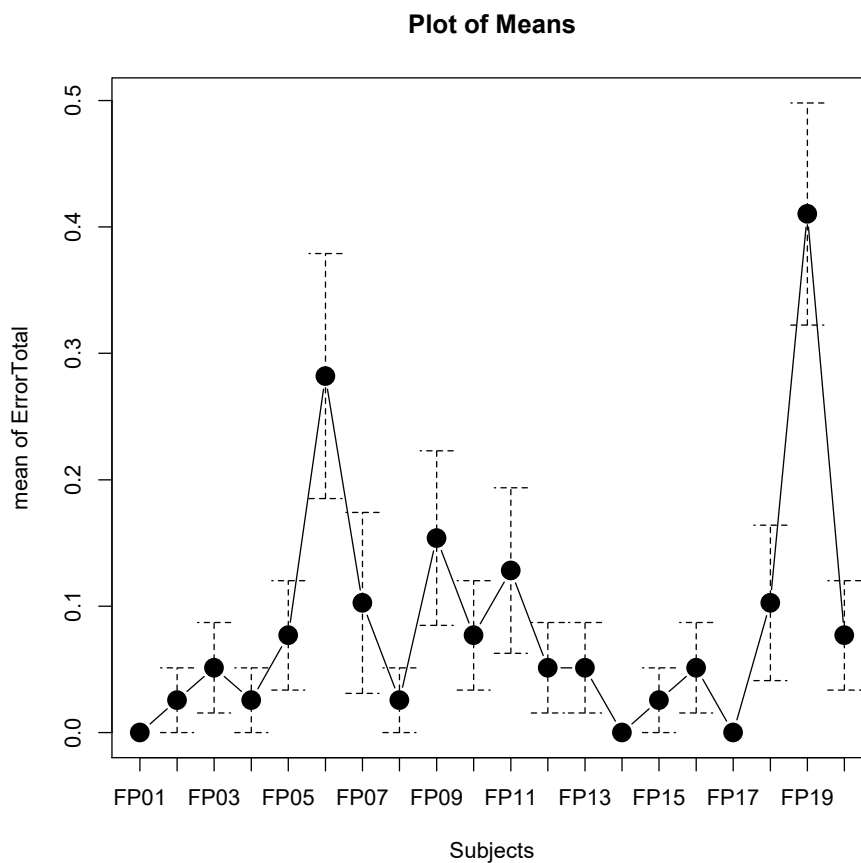


Figur 4.1: Plot of Means over forsøgspersonernes brug af de valgte keywords. Ud af x-aksen er forsøgspersonerne listet, og op ad y-aksen er gennemsnit af keywords brugt.

Ved at se på dette Plot of Means kan der ses, at forsøgsperson 11 ligger udenfor de andre

forsøgspersoners gennemsnit og tilhørende standardfejl, dermed har forsøgsperson 11 brugt signifikant færre keywords igennem hele forsøget end de andre forsøgspersoner. Det kan ses at nogle forsøgspersoner har benyttet sig mere af de forskellige keywords end andre forsøgspersoner, men selvom for eksempel forsøgsperson 1's gennemsnit og standardfejl ikke overlapper med forsøgsperson 6, så anses dette ikke for at være signifikant anderledes, da alle forsøgspersoners standardfejl, bortset fra forsøgsperson 11, overlapper med hinanden og dermed kan disse ses som værende ligeværdige.

Der kigges herefter på forsøgspersonernes fejl igennem forsøget, hvor der også laves et Plot of Means, hvilket kan ses på Figur 4.2.



Figur 4.2: Plot of Means over forsøgspersonernes fejl i løbet af forsøget. Ud af x-aksen er forsøgspersonerne listet, og op ad y-aksen er gennemsnit af fejl.

Der kan her ses, at forsøgsperson 19 ligger generelt højere end de andre forsøgspersoner, dog ligger forsøgsperson 6 også ret højt i forhold til de andre, hvilket gør, at forsøgsperson 19 vil blive set som ligeværdig med de andre forsøgspersoner. Fordi at det er forskellige forsøgspersoner, der enten er tæt på eller ligger som outliers i både brugen af keywords eller fejl i forsøget, så er der ingen forsøgspersoner, som vil blive taget ud af analysen.

Brug af keywords i de tre runder

Der vil nu blive undersøgt, om der er forskel på, hvor mange keywords der blev brugt i de tre forskellige runder. Dette gøres for at finde ud af, hvilken effekt det har haft at have initiativstyring og bekræftelse med i opgaverne. For at undersøge dette er det også nødvendigt at have med, hvilken rækkefølge runderne var præsenteret i, da dette kan have en indvirkning på resultaterne. Der laves derfor en lineær model, hvor der undersøges, om antallet af gange, forsøgspersonerne benytter sig af de valgte keywords, kan forklares ud fra, hvilken rækkefølge de to scenarier blev præsenteret i samt hvilken runde, de valgte keywords blev sagt i. Den lineære model kan ses på Figur 4.3.

```

                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                    4.173e-01  7.730e-02   5.398 9.01e-08 ***
Group[T.B]                     -2.487e-01  1.093e-01  -2.275 0.023181 *
Condition[T.B]                  2.550e-01  5.342e-02   4.773 2.17e-06 ***
Condition[T.C]                  4.083e-01  5.631e-02   7.251 1.02e-12 ***
Subjects[T.FP10]                -5.128e-02  9.877e-02  -0.519 0.603786
Subjects[T.FP11]                -5.641e-01  9.877e-02  -5.711 1.61e-08 ***
Subjects[T.FP12]                -2.564e-02  9.877e-02  -0.260 0.795250
Subjects[T.FP13]                -1.611e-15  9.877e-02   0.000 1.000000
Subjects[T.FP14]                -1.026e-01  9.877e-02  -1.038 0.299433
Subjects[T.FP15]                -1.538e-01  9.877e-02  -1.558 0.119759
Subjects[T.FP16]                -1.795e-01  9.877e-02  -1.817 0.069592 .
Subjects[T.FP17]                -2.051e-01  9.877e-02  -2.077 0.038164 *
Subjects[T.FP18]                -2.051e-01  9.877e-02  -2.077 0.038164 *
Subjects[T.FP19]                -3.846e-01  9.877e-02  -3.894 0.000107 ***
Subjects[T.FP2]                  1.538e-01  9.877e-02   1.558 0.119759
Subjects[T.FP20]                 7.692e-02  9.877e-02   0.779 0.436357
Subjects[T.FP3]                 -7.692e-02  9.877e-02  -0.779 0.436357
Subjects[T.FP4]                 7.692e-02  9.877e-02   0.779 0.436357
Subjects[T.FP5]                 -2.564e-02  9.877e-02  -0.260 0.795250
Subjects[T.FP6]                 -1.795e-01  9.877e-02  -1.817 0.069592 .
Subjects[T.FP7]                 -1.026e-01  9.877e-02  -1.038 0.299433
Subjects[T.FP8]                  NA          NA          NA          NA
Subjects[T.FP9]                 -4.103e-01  9.877e-02  -4.153 3.65e-05 ***
Group[T.B]:Condition[T.B]       2.600e-01  7.555e-02   3.442 0.000610 ***
Group[T.B]:Condition[T.C]       6.667e-02  7.963e-02   0.837 0.402770
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Figur 4.3: Resultatet af den lineære model over brugen af keywords, forklaret ud fra rækkefølgen og runderne.

Ved at se på variabelen “Order” (rækkefølge) viser p-værdien, at denne er signifikant, da p-værdien er under 0,05. Hvis der tages højde for type 1 fejl ved hjælp af Bonferroni’s formel (Field og Hole, 2003), så skal signifikansgrænsen for p-værdier divideres med antallet af tests, der udføres. I dette tilfælde skal de 0,05 divideres med 25, hvilket giver en ny signifikansgrænse på 0,002, hvilket rækkefølgens p-værdi ikke er under. Dermed tages rækkefølgen ud af den lineære model, og den udføres en gang til. Den nye lineære model kan ses på Figur 4.4.

Det kan ses på den lineære model, at der er signifikant forskel på, hvor mange keywords der bliver brugt imellem træningsrunden og de to andre runder, da der ved at gå fra interceptet, som er træningsrunden, til at gå over til de andre runder fremkommer meget lave p-værdier. For at visualisere forskellene og samtidig undersøge, om der er forskel på de to sidste runder, laves en anova test, hvor de tre runder bliver parvist sammenlignet. Disse sammenligninger kan ses på Figur 4.5.

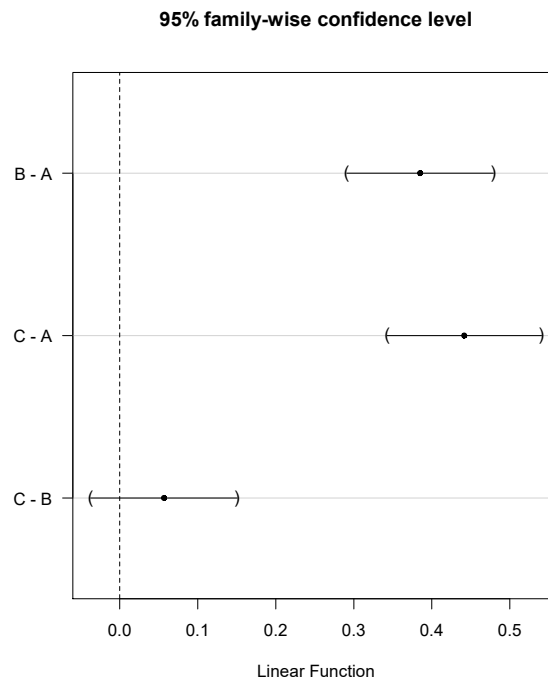
Der ses igen, at der er forskel, når træningsrunden bliver sammenlignet med de to andre runder. Der ses dog også, at der ikke er forskel på scenarierne om rejser og kalender, hvilket betyder, at bekræftelsen i rejsescenariet ikke øger brugen af keywords signifikant fra brugen af

```

Coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.571e-01  7.421e-02  4.811 1.81e-06 ***
Condition[T.B]  3.850e-01  3.805e-02 10.118 < 2e-16 ***
Condition[T.C]  4.417e-01  4.011e-02 11.012 < 2e-16 ***
Subjects[T.FP10] -1.795e-01  9.950e-02 -1.804 0.071644 .
Subjects[T.FP11] -5.641e-01  9.950e-02 -5.669 2.04e-08 ***
Subjects[T.FP12] -1.538e-01  9.950e-02 -1.546 0.122474
Subjects[T.FP13]  3.415e-15  9.950e-02  0.000 1.000000
Subjects[T.FP14] -2.308e-01  9.950e-02 -2.319 0.020644 *
Subjects[T.FP15] -1.538e-01  9.950e-02 -1.546 0.122474
Subjects[T.FP16] -3.077e-01  9.950e-02 -3.092 0.002058 **
Subjects[T.FP17] -2.051e-01  9.950e-02 -2.062 0.039586 *
Subjects[T.FP18] -3.333e-01  9.950e-02 -3.350 0.000848 ***
Subjects[T.FP19] -3.846e-01  9.950e-02 -3.866 0.000120 ***
Subjects[T.FP2]  2.564e-02  9.950e-02  0.258 0.796708
Subjects[T.FP20] -5.128e-02  9.950e-02 -0.515 0.606424
Subjects[T.FP3]  -7.692e-02  9.950e-02 -0.773 0.439704
Subjects[T.FP4]  -5.128e-02  9.950e-02 -0.515 0.606424
Subjects[T.FP5]  -2.564e-02  9.950e-02 -0.258 0.796708
Subjects[T.FP6]  -3.077e-01  9.950e-02 -3.092 0.002058 **
Subjects[T.FP7]  -1.026e-01  9.950e-02 -1.031 0.302963
Subjects[T.FP8]  -1.282e-01  9.950e-02 -1.289 0.197965
Subjects[T.FP9]  -4.103e-01  9.950e-02 -4.123 4.15e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Figur 4.4: Resultatet af den lineære model over brugen af keywords, forklaret ud fra runderne.

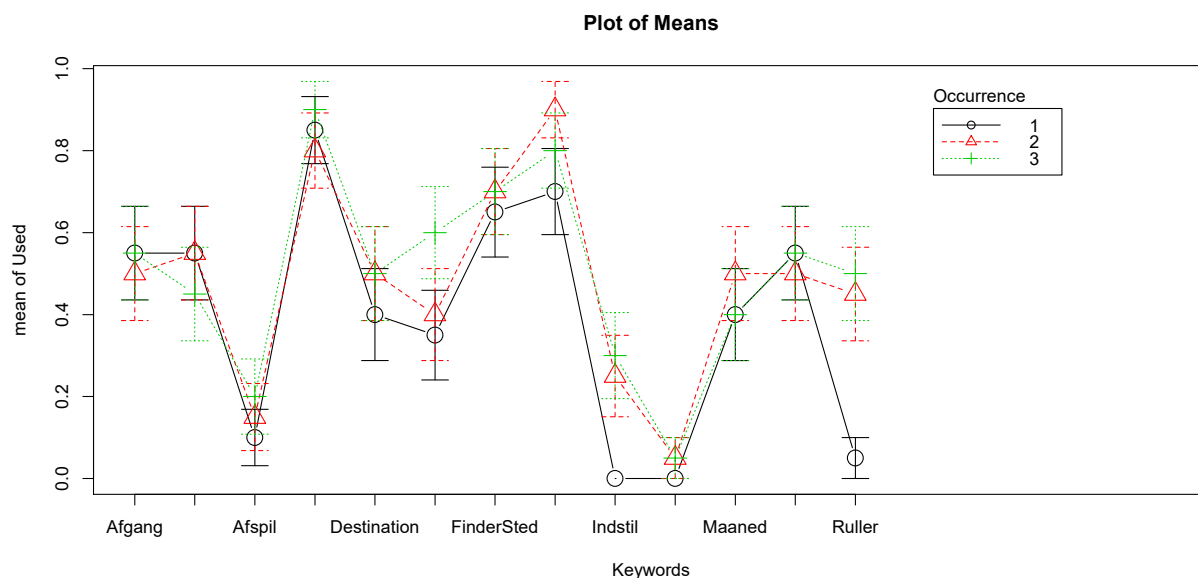


Figur 4.5: Parvis sammenligning af de tre runder. A er scenariet omhandlende det intelligente hjem, B er scenariet omhandlende bestilling af en rejse og C er scenariet omhandlende oprettelsen af en begivenhed i kalenderen.

keywords i kalenderscenariet. Derfor kan hypotesen, se Afsnit 3.7, om at nudging med bekræftelse og initiativstyring skulle være nemmere end kun med en af delene umiddelbart forkastes.

Brug af hvert enkelt keywords

Der vælges nu at se på, om antallet af gange der er benyttet de valgte keywords kan forklares ud fra, hvilke ord det er, og hvor om det er første, anden eller tredje forekomst af ordet. Dette gøres for at undersøge, om der er forskel på, hvilke ord der bruges hvor ofte. Der udføres en lineær model, hvor resultaterne for både ordene og forekomsterne signifikante, dog lå p-værdierne af forekomsterne på et niveau, hvor de med brug af Bonferroni's formel mod type 1 fejl ikke ville være signifikante længere. Dog vælges det at se på, om nogle keywords har været brugt mere efter første forekomst, hvilket er illustreret på Figur 4.6.



Figur 4.6: Plot of Means over brugen af de enkelte keywords fordel på de tre forekomster. Ud af x-aksen ligger de forskellige keywords, op ad y-aksen er gennemsnittet af brugen af de enkelte keywords, og 1, 2 og 3 er henholdsvis første, anden og tredje forekomst af det ord. Ordene på x-aksen står i denne rækkefølge: **Afgang, Afrejse, Afspil, Bagage, Destination, Emnet, Finder sted, Foregår, Indstil, Lamper, Måned, Påmind og Ruller.**

Nogle keywords har været brugt mere i anden og tredje forekomst, dog har dette hovedsageligt været de ord, som er at finde i træningsrunden, hvor der ikke er initiativstyring. Ved de resterende keywords ændres brugen ikke hen over tid, så derfor vælges det at tage forekomsterne ud af den lineære model, og derefter undersøge, hvilke keywords, som blev brugt flere gange end andre af forsøgspersonerne. Resultaterne af den lineære model kan ses på Figur 4.7.

Der ses ud fra den lineære model, at der er signifikante forskelle blandt ordene, derfor laves der et Plot of Means over brugen af de forskellige keywords, hvilket kan ses på Figur 4.8.

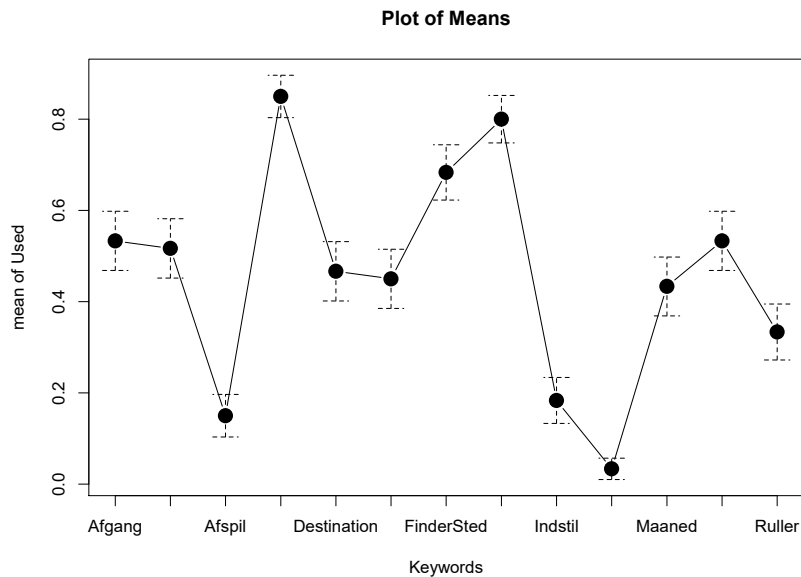
Ud fra Figur 4.8 ses det, at der er forskel på, hvor meget de forskellige keywords er blevet brugt i forsøget. De anvendte keywords kan inddeles i tre grupper alt efter, hvordan de ligger i plottet, en gruppe for de tre øverste, en gruppe for de tre nederste og en gruppe til de resterende keywords. De tre grupper kommer derfor til at se således ud:

```

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)      7.154e-01  8.526e-02  8.391 2.40e-16 ***
Keywords[T.Afrejse] -1.667e-02  7.685e-02  -0.217 0.828368
Keywords[T.Afspil] -3.833e-01  7.685e-02  -4.988 7.59e-07 ***
Keywords[T.Bagage]  3.167e-01  7.685e-02  4.121 4.20e-05 ***
Keywords[T.Destination] -6.667e-02  7.685e-02  -0.867 0.385955
Keywords[T.Emnet]  -8.333e-02  7.685e-02  -1.084 0.278556
Keywords[T.FinderSted]  1.500e-01  7.685e-02  1.952 0.051330 .
Keywords[T.Foregaar]  2.667e-01  7.685e-02  3.470 0.000550 ***
Keywords[T.Indstil] -3.500e-01  7.685e-02  -4.554 6.14e-06 ***
Keywords[T.Lamper]  -5.000e-01  7.685e-02  -6.506 1.41e-10 ***
Keywords[T.Maaned]  -1.000e-01  7.685e-02  -1.301 0.193581
Keywords[T.Paamind]  4.389e-16  7.685e-02  0.000 1.000000
Keywords[T.Ruller]  -2.000e-01  7.685e-02  -2.602 0.009439 **
Subjects[T.FP10]    -1.795e-01  9.532e-02  -1.883 0.060092 .
Subjects[T.FP11]    -5.641e-01  9.532e-02  -5.918 4.96e-09 ***
Subjects[T.FP12]    -1.538e-01  9.532e-02  -1.614 0.106955
Subjects[T.FP13]     4.029e-15  9.532e-02  0.000 1.000000
Subjects[T.FP14]    -2.308e-01  9.532e-02  -2.421 0.015717 *
Subjects[T.FP15]    -1.538e-01  9.532e-02  -1.614 0.106955
Subjects[T.FP16]    -3.077e-01  9.532e-02  -3.228 0.001301 **
Subjects[T.FP17]    -2.051e-01  9.532e-02  -2.152 0.031720 *
Subjects[T.FP18]    -3.333e-01  9.532e-02  -3.497 0.000498 ***
Subjects[T.FP19]    -3.846e-01  9.532e-02  -4.035 6.02e-05 ***
Subjects[T.FP2]     2.564e-02  9.532e-02  0.269 0.788007
Subjects[T.FP20]    -5.128e-02  9.532e-02  -0.538 0.590743
Subjects[T.FP3]     -7.692e-02  9.532e-02  -0.807 0.419931
Subjects[T.FP4]     -5.128e-02  9.532e-02  -0.538 0.590743
Subjects[T.FP5]     -2.564e-02  9.532e-02  -0.269 0.788007
Subjects[T.FP6]     -3.077e-01  9.532e-02  -3.228 0.001301 **
Subjects[T.FP7]     -1.026e-01  9.532e-02  -1.076 0.282282
Subjects[T.FP8]     -1.282e-01  9.532e-02  -1.345 0.179040
Subjects[T.FP9]     -4.103e-01  9.532e-02  -4.304 1.90e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Figur 4.7: Resultater af den lineære model over brugen af keywords, forklaret ud fra de enkelte keywords.



Figur 4.8: Plot of Means over brugen af de enkelte keywords. Ud af x-aksen ligger de forskellige keywords og op ad y-aksen er gennemsnittet af brugen af de enkelte keywords. Ordene på x-aksen står i denne rækkefølge: **Afgang, Afrejse, Afspil, Bagage, Destination, Emnet, Finder sted, Foregår, Indstil, Lamper, Måned, Påmind og Ruller.**

1. Gruppe

- Bagage
- Foregår
- Finder sted

2. Gruppe

- Påmind
- Afgang
- Afrejse
- Destination
- Emnet
- Måned
- Ruller

3. Gruppe

- Indstil
- Afspil
- Lamper

I den første gruppe er de tre keywords blevet anvendt signifikant flere gange end de andre, i den anden gruppe er de keywords blevet anvendt signifikant færre gange end i første gruppe men signifikant flere gange end anden gruppe og tredje gruppens keywords er anvendt signifikant færre gange end alle andre keywords. Fordi der er forskel blandt de tre grupper, og fordi der er signifikant forskel på, hvor meget nogle keywords er brugt frem for andre, så kan hypotesen, se Afsnit 3.7, omhandlende at nogle ord er nemmere at nudge end andre umiddelbart bekræftes.

4.2 Resultater af skalaerne

Forsøgspersonerne vurderede runderne på de fire skalaer, som beskrevet i Afsnit 3.5. Disse fire skalaer blev dog ikke valideret inden forsøget, hvilket gør, at forsøgspersonerne kan have forskellige forståelser for, hvad de egentlig vurderer og dermed ikke det samme som var tiltænkt med skalaerne. Dette gør, at dataet opsamlet fra skalaerne er mindre pålideligt, dog vælges det at analysere på dataet alligevel. Alt data fra skalavurderingerne kan findes på bilags-cd'en i mappen "Forsøg" under navnet "Skaladata".

Det vælges først at se, om der er forskel på, hvordan forsøgspersonerne har vurderet runderne alt efter, hvilken rækkefølge disse var præsenteret for forsøgspersonerne. Dette gøres ved at

sammenligne forsøgspersonernes skalavurderinger for hvert spørgsmål i hver runde, hvor vurderingerne er fordelt efter rækkefølgen. Der findes dog ingen forskel blandt spørgsmålene, når dataet er delt op efter rækkefølgen, og derfor antages det, at der ikke er nogen form for rækkefølgeeffekt.

Det vælges derefter at undersøge, om der er forskel på, hvordan det samme spørgsmål er vurderet blandt de forskellige runder. Dette gøres for at se, om en runde har påvirket forsøgspersonerne mere end en anden. Der udføres en række Paired-Samples Wilcoxon tests for hvert spørgsmål for at se, om der er forskel mellem nogle af runderne. Der findes her frem til to spørgsmål, hvor der er forskel på, hvordan forsøgspersonerne har vurderet dem.

Der er forskel på vurderingerne af det første skalaspørgsmål, som lød: "Hvor svært synes du det var at forstå opgaverne?", og forskellen var mellem scenarierne omhandlende rejser og kalender, hvilket kan ses på Figur 4.9.

```
Wilcoxon signed rank test with continuity correction
data: Calendar1 and Travell
V = 33, p-value = 0.03709
alternative hypothesis: true location shift is not equal to 0
```

Figur 4.9: Resultat af Paired-Samples Wilcoxon test ved første skalaspørgsmål mellem runderne med rejser og kalender.

Ved at se dataet igennem ligger forskellen i, at det har været sværere for forsøgspersonerne at forstå rejseopgaverne end at forstå kalenderopgaverne. Det andet spørgsmål, hvor der var forskel, var ved det tredje, som lød: "Hvor pålidelig virkede talesystemet til at være?", og forskellen var mellem scenarierne omhandlende hjemmestyring og kalender, hvilket kan ses på Figur 4.10.

```
Wilcoxon signed rank test with continuity correction
data: Calendar3 and Home3
V = 26, p-value = 0.03094
alternative hypothesis: true location shift is not equal to 0
```

Figur 4.10: Resultat af Paired-Samples Wilcoxon test ved tredje skalaspørgsmål mellem runderne med hjemmestyring og kalender.

Ved igen at se dataet igennem ligger forskellen her i, at talesystemet virkede mere pålideligt i hjemmestyringsopgaverne frem for i kalenderopgaverne.

4.3 Besvarelser fra exitinterviewet

Efter forsøget deltog alle forsøgspersoner i et exitinterview, hvor de blev spurgt ind til forskellige dele af forsøget. En række af resultaterne vil blive nævnt i dette afsnit. Alt data fra exitinterviewene kan findes på bilags-cd'en i mappen "Exitinterview" under forsøgspersonernes respektive numre.

Forsøgspersonerne blev spurgt ind til, om de lagde mærke til, om de ændrede deres ord til de samme som talesystemet brugte, hvortil de fleste af forsøgspersonerne sagde, at det gjorde de. Nogle mente at de ting, som talesystemet sagde gav mening i forhold til opgaven, andre følte det var nemmere at sætte et svar op til den ved at bruge spørgsmålet til at danne svaret med. Enkelte personer tænkte, at hvis de benyttede sig af de samme ord, så ville talesystemet have nemmere ved at forstå dem. Der blev dog også nævnt et par tilfælde, hvor nogle af de ord, som talesystemet brugte, ikke gav mening for forsøgspersonerne, og de fortsatte dermed med at benytte deres egne ord. Fire af forsøgspersonerne begyndte på forskellige tidspunkter at afprøve talesystemet ved at benytte sig af andre ord og sætninger, for at se, i hvilken grad talesystemet kunne forstå dem.

Til spørgsmålene omkring den valgte stemme til forsøget gav de fleste af forsøgspersonerne udtryk for, at den enten var behagelig, neutral eller venlig. Der var delte meninger om, hvorvidt stemmen virkede som et menneske eller som en maskine, men ingen forsøgspersoner gav udtryk for, at stemmen virkede ubehagelig for dem. Fem af forsøgspersonerne gav til kende, at de synes, at stemmen virkede bekendt for dem, hvilket gav dem mere tryghed, da den virkede familiær. En mente at det kunne være en stemmeskuespiller, som han havde hørt før, og en anden gættede på, at det var ham, som lagde stemme til vortesvinet Pumba fra Løvernes Konge, hvilket er den persons stemme, som var valgt til forsøget.

Vedrørende scenarie C omhandlende kalenderen var mange af forsøgspersonerne enige om, at der var noget underligt ved den måde, som talesystemet sagde "Okay" på. De fleste ville gerne have haft feedback på deres opgaver, så de kunne blive mere sikre på, om de havde udført opgaverne korrekt, mens andre synes at systemet virkede mindre pålideligt, irriterende eller direkte flabet ved blot at svare "Okay".

Forsøgspersonerne skulle også svare på, hvilken af de tre runder, som de bedst kunne lide. Halvdelen af forsøgspersonerne mente, at træningsrunden var den bedste, mest fordi emnet virkede meget brugbart for forsøgspersonerne, da de kunne se anvendelsen i deres egne hjem. Seks af forsøgspersonerne mente, at scenarie B med rejsebestillingen virkede bedst, fordi der var bekræftelse på, og fordi de blev guidet igennem det af systemet, og to forsøgspersoner mente, at scenarie C virkede bedst. Enkelte mente også at det var den runde, de var igennem sidst, fordi de følte sig mere trygge med systemet.

To af forsøgspersonerne lavede markant flere fejl end andre i forsøget, hvilket kom til udtryk i exitinterviewene. Forsøgsperson 6 følte at det var unaturligt at skulle lave lange og kunstige sætninger og ville gerne kunne give hele opgaven i én sætning. Han følte at han havde svært ved at regne ud, hvad han skulle i de forskellige opgaver, og han manglede som mange andre også bekræftelse i runden med kalenderopgaverne. Han tænkte at han kunne have klaret sig bedre, hvis der havde været en GUI indblandet og han ville samtidig hellere have haft en kvindestemme til systemet, da han stolede mere på Siri end på talesystemet i forsøget. Han synes også, at stemmen lød deprimerende, og den mindede ham om Microsoft Sam. Forsøgsperson 19 ville også gerne have givet kortere svar til talesystemet, især fordi talesystemet lagde op til, at der kunne komme korte svar tilbage. Han blev irriteret på de sætninger, som stemmen sagde til sidst, da han havde hørt den samme sætning signifikant flere gange end de fleste andre forsøgspersoner, og han mente også, at stemmen mindede ham om Microsoft Sam.

Sidst i forsøget blev forsøgspersonerne spurgt om, hvorvidt de havde tænkt på noget enten under forsøget eller i løbet af exitinterviewet. Her gav fem forsøgspersoner udtryk for, at de var interesserede i at vide, hvorvidt der var faktisk var et system foran dem, eller om de blev snydt

af forsøgslederne.

DISKUSSION

I dette kapitel vil der blive diskuteret, hvorfor dataet ser ud som det gør, og set på hvilke elementer der kunne have haft en indflydelse, i forhold til hvordan forsøget er udført, og hvad forsøgspersonerne har givet udtryk for i exitinterviewene.

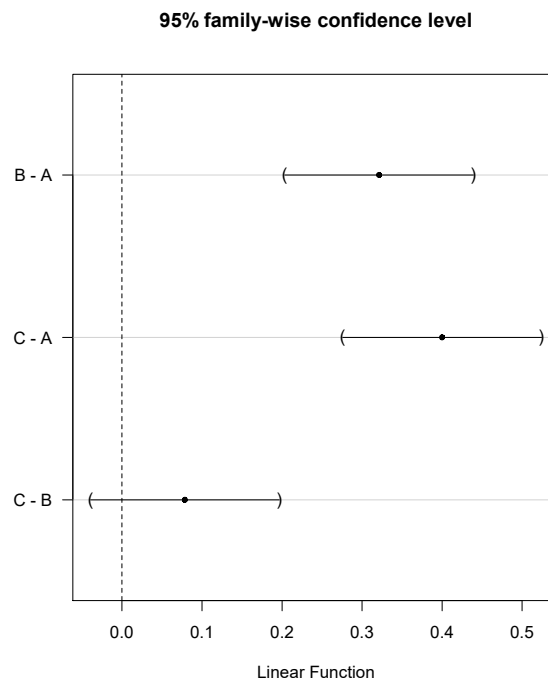
5.1 Forsøgspersonernes anvendelse af keywords

Forsøgsperson 11 er den eneste, som anvendte keywordene signifikant lavere end alle andre, hvilket kan ses på Figur 4.1. I exitinterviewet gav han udtryk for at han ikke er så kreativ med at finde på sætninger, så det kan antages at han benyttede de samme slags sætninger, som tidligere i forsøget havde virket for ham. Ses der på forsøgsperson 11's antal fejl, hvilket kan ses på Figur 4.2, ligger han som en del af gruppen, hvilket er grunden til at han ikke bliver anset som en outlier. Det kan dog være at plot of means of totale anvendte keywords ville have ændret sig, hvis han var blevet betragtet som en outlier, og var taget ud af dette plot.

5.2 Forsøgsrundernes indflydelse

Når der ses på hvorvidt om forsøgspersonerne er blevet nudget lige meget i de forskellige forsøgsrunder, viser det at forsøgspersonerne har anvendt signifikant flere keywords i scenarie B og scenarie C, end i træningsrunden. Dette kan skyldes måden, hvorpå de forskellige runders talestruktur er opstillet. Som beskrevet i 3 så bekræfter træningsrunden kun, scenarie C initiativstyrer kun, og scenarie B gør begge dele. Det kan derfor antages at en initiativstyret talestruktur påvirker forsøgspersonerne mere til at anvende keywords end en talestruktur hvor talesyste-

met udelukkende bekræfter forsøgspersoners input. Dette kan skyldes, at ved denne form for talestruktur hører forsøgspersonerne ikke keywordet i opgave 1, før efter de har kommet med deres input til talesystemet. De kender derfor ikke keywordene før i opgave 2 og 3. Er det så fair at sammenligne runderne på denne måde? Måske ikke, derfor er der opstillet en anova parvis sammenligning, hvor opgave 1 for alle runderne ikke er inkluderet, som kan ses på Figur 5.1.



Figur 5.1: Parvis sammenligning af opgave 2 og 3 i de tre runder. A er scenariet omhandlende det intelligente hjem, B er scenariet omhandlende bestilling af en rejse og C er scenariet omhandlende oprettelsen af en begivenhed i kalenderen.

Denne figur viser dog, at der stadigvæk er en signifikant forskel på træningsrunden når den sammenlignes med scenarie B og C. Dette er dog en træningsrunde, som er lavet for at gøre forsøgspersonerne fortrolige med opgaven og interaktionen med talesystemet, hvilket kan betyde at forsøgspersonen havde brug for en tilvendingsperiode, som kan have haft en indflydelse på deres måde at respondere på. Dog hvis der ses på forsøgspersonernes vurdering på skalaerne, og deres besvarelser i exitinterviewet, kan det antages, at denne tilvendingsperiode ikke har haft indflydelse på deres brug af keywords.

Endvidere er der stadigvæk ingen signifikant forskel på anvendelsen af keywords i scenarie B og C, hvilket kan betyde, at bekræftelse ikke har givet nogen yderligere effekt, når det er sammensat med initiativstyring. Ud fra denne observation kan den opstillet hypotese derfor stadigvæk forkastes, idet at der ikke er forskel mellem hvorvidt der anvendes initiativstyring og bekræftelse, eller kun initiativstyring i talesystemet.

5.3 Forekomst af keywords

I afsnit 4.1 blev der redegjort for, at nogle ord blev anvendt mere igennem forsøget af forsøgspersonerne end andre. Ud fra Figur 4.8 blev der identificeret tre grupper af ord, som er inddelt efter hvor meget de er blevet brugt af forsøgspersonerne. Hvis ordene betragtes i de opstillede grupper, er der umiddelbart tre keywords, som er signifikant mere anvendt end de resterende (**Bagage**, **Foregår**, **Finder sted**), og tre ord som er signifikant mindre anvendt end de resterende (**Indstil**, **Afspil**, **Lamper**), hvilket kan tyde på, at der er en forskel på graden af nudging for de forskellige ord. Dette kan skyldes ordenes gængse brug i dagligdagen i de sammenhænge, som de er sat i i de forskellige scenarier, samt den talestruktur der er anvendt af talesystemet i de respektive scenarier. Forsøgspersonerne gav også selv udtryk for i exitinterviewene, at ord som for eksempel **Lamper** ikke er noget, som de normalt ville bruge i den sammenhæng de blev præsenteret i, hvorimod at et ord som **Indstil**, giver mere mening at anvende i den sammenhæng den blev præsenteret. De tre ord, som er de mindst anvendte, forekommer desuden alle i træningsrunden, hvor der jo kun anvendes bekræftelse til at nudge forsøgspersoner. Det kan derfor diskuteres hvorvidt dette har haft en effekt på brugen af ordene, da det allerede er observeret at brugen af keywords i træningsrunden er signifikant mindre end i de to andre runder. Derudover kan det eksempel de blev præsenteret for i introduktionen have influeret forsøgspersonernes måde at danne sætningerne på, således at de blev nudget til at forme sætninger ud fra den standard, da talesystemet ikke har præsenteret dem for andre alternativer.

Endvidere kan det ses ud fra Figur 4.6, at nogle af ordene var blevet anvendt signifikant færre gange i opgave 1 i den respektive runde end i opgave 2 og 3 i samme runde. Disse ord var henholdsvis **Indstil** og **Ruller**. Begge ord fremtræder i træningsrunden, og det kan derfor tyde på, at forsøgspersonerne var blevet nudget til at anvende disse ord til opgave 2 og 3, på grund af den bekræftelse de fik i opgave 1. Dog er ordene **Afspil** og **Lamper** ikke blevet nudget nok, til at forsøgspersonerne er begyndt at anvende de keywords, som talesystemet har brugt.

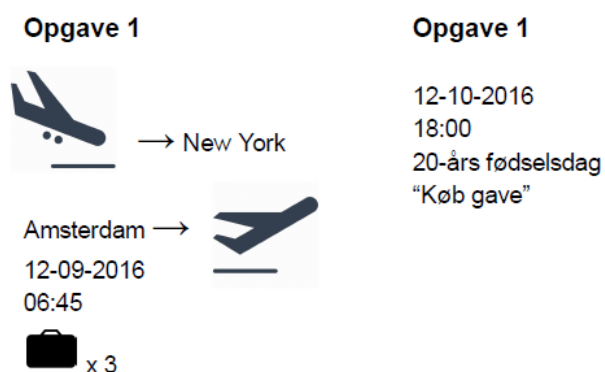
Ud fra disse observationer kan det antyde, at det ikke er ligegyldigt, hvilke ord der vælges at anvende når der designes et talesystem, da det ville være fordelagtigt at tilgodese brugerens ordbrug.

5.4 Skalaer på tværs af runder

Under analysen af skalaerne blev der ved hjælp af paired samples wilcoxon tests kun fundet frem til to områder, hvorpå der har været en forskel mellem runderne. Disse to områder er henholdsvis: sværhedsgraden mellem scenarie B og C, og pålideligheden af talesystemet i træningsrunden og scenarie C.

I forhold til sværhedsgraden af opgaverne, har forsøgspersonerne ifølge Figur 4.9 vurderet scenarie B's opgaver til at være sværere at forstå end scenarie C. Dette kan skyldes at der i scenarie B er anvendt ikoner til at beskrive nogle af delopgaverne, hvorimod scenarie C's delopgaver udelukkende består af tekst (se Figur 5.2). Det kan derfor antages at ikonerne i scenarie B har sværere at tolke end teksten i scenarie C. Dog blev der nævnt i exitinterviewene at rækkeføl-

gen hvori de skulle bestille en rejse i scenarie B var med til at forvirre dem, da de hellere ville bestemme deres afgang før destination, hvor de blev bedt om det modsatte af talesystemet i forsøget.



Figur 5.2: Figuren viser et eksempel på hvordan opgaverne var opstillet i henholdsvis scenarie B (til venstre) og scenarie C (til højre).

Derudover vurderede forsøgspersonerne desuden, at talesystemet i træningsrunden virkede mere pålideligt end i scenarie C (se Figur 4.10). Her kan det have haft indflydelse, at træningsrunden besvarede forsøgspersonerne med en bekræftelse, som gentagede det de sagde, hvilket talesystemet ikke gjorde i scenarie C, men derimod kun med "Okay". Dette blev også nævnt i exitinterviewene, hvor nogle gav udtryk for at det var rart at den bekræftede det de sagde, så de var sikre på at de havde løst opgaven korrekt. Derimod var der en enkelt, som sagde at det ligefrem virkede mindre pålideligt at systemet kun besvarede med "Okay". Dog er scenarie B ikke vurderet hverken mere eller mindre pålideligt end de to andre runder, på trods af at denne havde samme form for feedback, som træningsrunden havde. Der kan derfor stilles spørgsmål om, om det alene er bekræftelsen der har gjort det, eller om der er andet, som kan have gjort enten træningsrunden mere pålidelig, eller scenarie B mindre pålidelig. Skalaerne var desværre ikke valideret før forsøget, så det er svært at sige om dette i sig selv ikke kan have haft en indflydelse på dataet.

5.5 Exitinterviewene

Som nævnt i Afsnit 4.3 gav de fleste af forsøgspersonerne udtryk for, at de lod sig påvirke af, hvad talesystemet sagde. Dette var af forskellige grunde såsom at talesystemet bedre ville kunne forstå forsøgspersonernes kommandoer, det var nemmere at sige det samme, eller at de ord, som talesystemet brugte virkede passende til opgaverne. Dette tyder umiddelbart på, at det er lykkedes at nudge forsøgspersonerne til at benytte de samme ord som talesystemet.

I forhold til valget af stemme var de fleste forsøgspersoner positive. Det at stemmen har været set som en venlig om behagelig stemme har formodentlig gjort, at forsøgspersonerne har følt sig mere tilpas til forsøget. Der var dog enkelte forsøgspersoner, som genkendte stemmen i større eller mindre grad, hvorved de kan have overført deres relation fra stemmen over til ta-

lesystemet. Hvis forsøgspersonerne havde haft dårlige relationer til den stemme kan det have påvirket forsøget, dog gav ingen af forsøgspersonerne udtryk for det, hvilket kunne betyde, at den valgte stemme var et fornuftigt valg.

Problemet med ordet “Okay” opstod allerede under pilotforsøget, hvor der blev lavet en ny optagelse, hvilket blev vurderet til at være mere positiv end den første. Dog virker det til, at ordet skulle have været endnu mere positivt, da mange af forsøgspersonerne synes, at det lød mærkelig. Det blev dog samtidig nævnt med, at de gerne ville have haft en bekræftelse i stedet for, så det, at bekræftelsen manglede i det scenarie kan have gjort, at det ord blev opfattet mindre positivt.

De to forsøgspersoner, som lavede flere fejl end de resterende, gav begge udtryk for, at talesystemet gav dem associationer til Microsoft Sam, og måden de sagde dette på hentydede til, at de ikke var begejstrede for denne. Det virker derfor til, at fordi de havde en knap så god oplevelse med talesystemet, så overførte de deres associationer fra et tidligere tekst-til-tale-system til talesystemet. Det kunne derfor tyde på, at hvis talesystemet giver forsøgspersoner en dårlig oplevelse, så vil de sammenligne det med lignende systemer og muligvis knytte dem sammen, og forsøgspersonerne vil formentlig have svært ved at komme til at tro på talesystemet igen.

KONKLUSION

Dette kandidatspeciale tager udgangspunkt i, hvorvidt brugere kan nudges udelukkende auditivt til at anvende specifikke ord, når de interagerer med et talesystem. Der er kort blevet redegjort for udviklingen af talesystemer og problemer, som kan opstå med disse. Der er inddraget relevante teorier, der ligger til grund for dette projekt. I denne forbindelse er der blevet undersøgt, hvordan tre forskellige talestrukturer i et talesystem har påvirket en række forsøgspersoner til at anvende udvalgte keywords.

Forsøget bestod af at forsøgspersoner skulle interagere med et talesystem, for at udføre en række opgaver til tre forskellige scenarier, hvori at talesystemet anvender specifikke keywords, når denne kommunikerer med forsøgspersonerne, i et forsøg på at nudge dem til at anvende disse selv samme keywords. Forsøget skulle besvare hypoteserne som lød på at det er nemmere at nudge med både en initiativstyrende og bekræftende talestruktur i et talesystem, frem for anvendelsen af kun en af disse, og hypotesen omhandlende at visse er nemmere at nudge end andre.

Ud fra resultaterne fra forsøget kan den første hypotese forkastes, da dataet tyder på, at initiativstyring uden bekræftelse har nudget forsøgspersonerne lige så meget, som hvis begge var anvendt på samme tid. Derudover kan den anden hypotese bekræftes, da dataet tyder på, at der er signifikant forskel på, hvor ofte visse ord bliver anvendt i de opstillede scenarier.

Ses der på skalavurderinger af de forskellige runder, er der en tendens til, at bekræftelse giver en signifikant højere pålidelighed, end hvis bekræftelse udelades af talesystemet. Dog er skalerne ikke valideret før anvendelse, hvilket gør at denne observation ikke kan konkluderes med sikkerhed. I exitinterviewene gav flere forsøgspersoner også udtryk for, at visse ord, som talesystemet anvendte, gav bedre mening at anvende i opgaverne, end andre ord som talesystemet anvendte, hvilket understøtter den anden hypotese.

Tages alt det ovenstående i betragtning, så kunne det tyde på at brugere til en vis grad kan nudges til at anvende specifikke ord til interaktion med et talesystem. Dog er det ikke ligegyldigt, hvilke ord der anvendes i visse situationer, da brugere ikke lader sig nudge af ord, som ligger dem for fjernt fra konteksten.

PERSPEKTIVERING

I dette kapitel vil der blive set på hvordan resultaterne fra undersøgelsen udført i denne rapport kan anvendes som råd eller guidelines til fremtidig design af talesystemer, og hvordan undersøgelse kunne udbygges for at skabe mere validt data, eller tage det i en anden retning.

7.1 Videre design af talesystemer

Ud fra resultaterne fra undersøgelsen blev der konkluderet at effekten af nudging er afhængig af ordet i den bestemte kontekst det anvendes. Dette vil sige, at det ikke er fordelagtig at anvende hvilket som helst ord eller udtryk til hvilken som helst situation. Derfor er det vigtigt at et fremtidigt talesystem anvender nogle bestemte ord alt efter hvilken situation eller kontekst det ønsker input under, også selvom det måske praktisk set giver mening. Et eksempel er udtrykket **tænd lamperne** i træningsrunden af forsøget. Dette keyword blev kun anvendt af de færreste forsøgspersoner, som i stedet valgte at bruge et udtryk som **tænd lyset**. Forsøgspersonerne gav nemlig udtryk for i exitinterviewene, at selvom tænd lamperne i princippet giver mening, så er det ikke et udtryk de kunne finde på at anvende i den sammenhæng. Derfor er dette vigtigt at tage i betragtning når et talesystems ordvalg skal bestemmes til specifikke situationer.

Derudover var der også en tendens til at initiativstyringen nudgede forsøgspersoner mere end når der kun blev anvendt bekræftelse, men mange af forsøgspersonerne sagde dog at de efter lidt tid med talesystemet ikke brød sig om at skulle gennemgå hvert enkelt trin af de forskellige opgaver, men at de gad godt at de kunne inputte mere information på én gang. Dette vil dog fjerne en del af initiativstyringen, som jo var den del der nudgede mest. Det kan derfor overvejes hvor stor en del af initiativstyringen der burde fjernes for både at optimere brugernes interaktion med talesystemet, men også for at sørge for at nudgingen af ord stadigvæk sker. Endvidere er det en god idé at have en form gentagende bekræftelse af hvilket input en bruger har givet

til et talesystem, især hvis der ikke er en GUI, da det er betryggende for brugeren, at han eller hun ved, hvad det er, talesystemet har modtaget af input. Det vil derfor være hensigtsmæssigt at undgå responser som, "Okay", da det kun siger at talesystemet har modtaget et input, men det siger ikke noget om hvad for et input, hvis ikke der er en form for direkte visuel feedback.

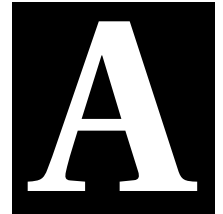
7.2 Videreundersøgelse

Til forsøget blev der anvendt en række ord, der fungerede som de keywords, der blev forsøgt nudget igennem forsøget. Disse keywords var dog alle nogle, som blev udvalgt ud fra antagelser omkring visse scenarier og situationer, uden nogen reel validering af hvorvidt om disse antagelser egentlig stemte overens med hvad virkeligheden måske siger. Derfor kunne det være fordelagtig, hvis et lignende forsøg skulle opstilles, at undersøge hvilke ord der burde anvendes i sådan et forsøg. Dette kunne for eksempel være gennem en word elicitation undersøgelse, hvor ordene kan blive defineret og valideret til en given situation eller scenarie af flere respondenter, for at give et bedre grundlag for udvalget af ord et talesystem skal anvende.

Endvidere kunne det undersøges hvorvidt talesystemets stemme har en effekt på om brugeren bliver nudget. En lignende undersøgelse, som er udarbejdet i denne rapport, kunne opsættes, men med forskellige stemmer. Her kunne der ligges vægt på om der er forskel på om det er en mand, en kvinde eller om talesystemet slet ikke lyder menneskelig, men lyder mere som en robot. Nogle forsøgspersoner gav i exitinterviewene udtryk for at den stemme som blev anvendt i forsøget lød bekendt, hvilket gjorde at systemet kom til at virke mere pålideligt, da det af den grund virkede mere familiært.

LITTERATUR

- A. Field og G. Hole. *How to Design and Report Experiments*. SAGE Publications, 2003. ISBN 9780761973836. 26
- W. W. Gaver. Technology Affordances. I *Proceedings of the ACM CHI 91 Human Factors in Computing Systems Conference*, side 79–84. ACM, jun 1991. 10
- J. J. Gibson. *The Ecological Approach to Visual Perception*. Lawrence Erlbaum, 1986. ISBN 978-89859-959-6. 9
- T. Hempel, editor. *Usability of Speech Dialog Systems: Listening to the Target Audience*. Springer Publishing Company, Incorporated, 2010. ISBN 3642097022, 9783642097027. URL <http://link.springer.com/book/10.1007/978-3-540-78343-5>. 11
- E. Luger og A. Sellen. “Like Having a Really bad PA”: The Gulf between User Expectation and Experience of Conversational Agents. I *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, side 5286–5297. SIGCHI, ACM, may 2016. ISBN 978-1-4503-3362-7. 8
- M. F. McTear. *Spoken Dialogue Technology: Toward the conventional user interface*. Springer, 2004. ISBN 1-85233-672-2. 7, 8
- M. Pinola. History of voice recognition: from Audrey to Siri, 2011. URL <http://www.itbusiness.ca/news/history-of-voice-recognition-from-audrey-to-siri/15008>. 7
- B. Reeves og C. Nass. *The Media Equation: How people treat computers, television, and new media like real people*. CSLI Publications, 1st udgave, 1998. ISBN 1-57586-053-8. 8



FORSØGSDOKUMENTER

A.1 Materialeliste

Dette bilag indeholder information omkring det materiale, som er anvendt til forsøget.

Til forsøget blev der anvendt dette fysiske materiale:

- Pioneer Stereo Amplifier A-616
- Asus U36S Bærbar
- En højttaler
- Ericsson Centrum GmbH Hannover samtaleanlæg

Derudover er der blevet anvendt en række ikoner til forsøgsopgaverne. En liste over de forskellige links til siderne ikonerne kommer fra kan ses i nedenstående punkter:

- Træningsrunden
 - <http://simpleicon.com/wp-content/uploads/light-bulb-3.png>
 - http://thumb10.shutterstock.com/photos/thumb_large/384199/108092945.jpg
 - <http://www.iconsdb.com/icons/preview/black/temperature-2-xxl.png>
 - https://image.freepik.com/free-icon/livingroom-window-with-curtains_318-63411.jpg
- Scenarie B
 - https://cdn3.iconfinder.com/data/icons/airport/80/Airport_icons-02-512.png
 - https://cdn3.iconfinder.com/data/icons/airport/80/Airport_icons-03-512.png
 - <http://downloadicons.net/sites/default/files/briefcase-icon-46075.png>

A.2 Forsøgsprogram

Dette bilag viser det forsøgsprogram som forsøgslederen brugte til at styre talesystemet med, og hvordan det er opstillet. Forsøgsprogrammet er lavet af en række lydfiler, som er linket til forskellige knapper i et Powerpoint, og det er udelukket designet for at hjælpe forsøgslederen til at kunne styre talesystemet således det passer med den tiltænkte gennemgang. Det egentlige forsøgsprogram kan findes på bilags-cd'en under navnet "Talesystemet".

Træningsrunde
Opgave1

Forstår ikke



Hele sætninger



Husk skala



Introduktion



Tænd i køkkenet

Lamperne



24 grader

Indstiller temp.



Rockmusik

Afspiller musik



Rul fra i soveværelse

Gardiner



Dette bilag indeholder introduktionen og samtykkeerklæringen, som forsøgspersonerne blev givet til forsøget. Derefter følger spørgsmålene, som forsøgspersonerne blev stillet i exitinterviewet.

A.3 Introduktion

Velkommen til!

I dette forsøg vil du blive bedt om at udføre en række opgaver ved at interagere med et talesystem, hvor vi gerne vil have lov til at filme dig imens. Forsøget består af tre runder - en træningsrunde og to forsøgsrunder - som hver består af tre opgaver i et bestemt scenarie.

Talesystemet er kun en prototype, så det er begrænset, hvad du kan sige til det, så derfor beder vi dig følge opgaverne, som du kan finde i dette kompendium. Opgaverne består i, at du skal få talesystemet til at udføre nogle funktioner ved at tale til det. Når du har udført en opgave, og du er klar til næste opgave, bedes du sige: **Klar**, så næste opgave kan begynde.

Når du er færdig med alle opgaverne i en runde, skal du vurdere din oplevelse med talesystemet på nogle skalaer, som du kan finde på papirerne ved siden af dig.

I løbet af din interaktion med talesystemet vil vi gerne have, at du benytter dig af hele sætninger og helst undgår korte svar. Med en hel sætning mener vi ligesom i dette eksempel:

Talesystem: "Hvad vil du gerne have at spise?" Du svarer: "Jeg vil gerne have pizza"	Talesystem: "Hvad vil du gerne have at spise?" Du svarer: "Pizza"
Dette vil talesystemet forstå.	Dette vil talesystemet <u>ikke</u> forstå.

Hvis du har spørgsmål til forsøgslederne under forsøget, så har du mulighed for at spørge til hver en tid ved hjælp af samtaleanlægget, der står på bordet. Vi vil gøre dig opmærksom på, at vi tester talesystemet og ikke dig.

Når forsøget er forbi, vil vi meget gerne have dig til at deltage i et interview, hvor vi vil snakke om forsøget. Vi gør dig opmærksom på, at du til hver en tid kan afbryde forsøget. Når du har underskrevet samtykkeerklæringen, og du ikke har flere spørgsmål, er vi klar til at påbegynde træningsrunden.

A.4 Samtykkeerklæring

Dette forsøg er opsat af to studerende fra 10. semester på Produkt- og Designpsykologi ved Aalborg Universitet. Som deltager vil du være anonym og du vil i projektrapporten fremgå ved et nummer.

Med underskrift på dette dokument giver du tilladelse til at blive videooptaget under forsøget. Samtidig giver du tilladelse til, at data fra forsøget må anvendes i forbindelse med udarbejdelse af et kandidatspeciale. Du kan til hver en tid afbryde forsøget, og alt data opsamlet om dig vil da blive slettet.

Dato/Underskrift

A.5 Exitinterview

Forsøgspersonerne var som en del af forsøget med i et exitinterview. Spørgsmålene, som forsøgslederen stillede til forsøgspersonerne, kan ses herunder.

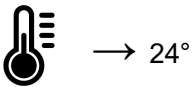
- Hvordan synes du, at det gik?
- Hvordan synes du, opgaverne var?
- Har du erfaringer med talesystemer inden idag?
 - Hvis ja, hvilke og i hvor høj grad?
- Hvad synes du om talesystemet?
- Hvad synes du om at snakke til talesystemet?
- Hvad synes du om stemmen, vi brugte til talesystemet?
 - Kom du til at tænke på nogen eller noget, da du hørte stemmen?
 - Synes du, at stemmen lød bekendt?
 - Hvilken effekt havde det, at du kendte/ikke kendte stemmen?
- Hvad synes du om måderne talesystemet kommunikerede på igennem de forskellige runder?
 - Hvilken måde/runde foretræk du?
- Lagde du mærke til om de ord du brugte til funktionerne blev påvirket af talesystemet?
- Hvad synes du om måden talesystemet så ud på?
 - Påvirkede det din oplevelse med det?
- Har du nogen yderligere kommentarer?
- Køn
- Alder
- Beskæftigelse (studieretning)

A.6 Opgaverne til scenarie A - Træningsrunden

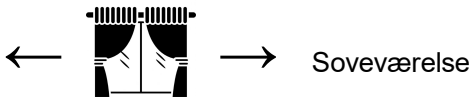
Træningsrunde

I dette scenarie skal du interagere med et "intelligent hjem". De funktioner du skal kontrollere er, belysning, temperatur, musik og gardiner, i den rækkefølge. Vi vil gerne have, at du benytter dig af hele sætninger og helst undgår korte svar.

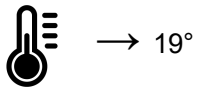
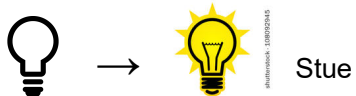
Opgave 1



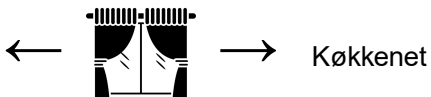
Musik -> Rock



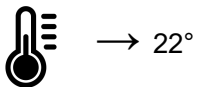
Opgave 2



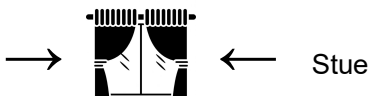
Musik -> Hip Hop



Opgave 3



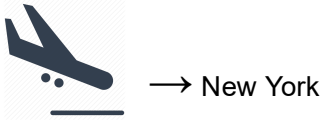
Musik -> Jazz



A.7 Opgaverne til scenarie B

I dette scenarie skal du bestille en flyrejse. Dette skal inkludere, hvor du skal hen, hvor du rejser fra, dato, tidspunkt og hvad du tager med, i den rækkefølge. Vi vil gerne have, at du benytter dig af hele sætninger og helst undgår korte svar.

Opgave 1



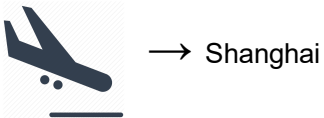
Amsterdam →

12-09-2016

06:45



Opgave 2



Frankfurt →

25-09-2016

12:30



Opgave 3



Aalborg →

02-10-2016

17:00



A.8 Opgaverne til scenarie C

I dette scenarie skal du lave en begivenhed i din kalender. Det skal inkludere dato, tidspunkt, hvad det er, og en ekstra notifikation, i den rækkefølge. Vi vil gerne have, at du benytter dig af hele sætninger og helst undgår korte svar.

Opgave 1

12-10-2016
18:00
20-års fødselsdag
"Køb gave"

Opgave 2

07-12-2016
13:45
Bryllup
"Stryg tøj"

Opgave 3

27-05-2017
10:00
Karneval
"Find kostume"

A.9 Talespor til scenarie A - Træningsrunden

Træningsrunde

Opgave 1

TS: Velkommen til dit stemmestyrede intelligente hjem. Hvad kan jeg hjælpe dig med?

FP:

- Tænd lamperne i køkkenet
- Tænd lyset i køkkenet

TS: Tænder lamperne i køkkenet.

FP:

- Indstil temperaturen til 24 grader.
- Sæt temperaturen til 24 grader.
- Ændre temperaturen til 24 grader.
- Temperaturen skal ændres til 24 grader.

TS: Indstiller temperaturen til 24 grader

FP:

- Afspil noget rockmusik
- Spil noget rockmusik
- Sæt noget rockmusik på

TS: Afspiller rockmusik

FP:

- Rul gardinerne fra i soveværelset
- Træk gardinerne fra i soveværelset

TS: Ruller gardinerne fra i soveværelset. Du er nu færdig med opgave 1.

Opgave 2

TS: Velkommen til dit stemmestyrede intelligente hjem. Hvad kan jeg hjælpe dig med?

FP:

- Tænd lamperne i stuen
- Tænd lyset i stuen

TS: Tænder lamperne i stuen.

FP:

- Indstil temperaturen til 19 grader.
- Sæt temperaturen til 19 grader.
- Ændre temperaturen til 19 grader.

- Temperaturen skal ændres til 19 grader.

TS: Indstiller temperaturen til 19 grader

FP:

- Afspil noget hiphopmusik
- Spil noget hiphopmusik
- Sæt hiphopmusik på

TS: Afspiller hiphopmusik

FP:

- Rul gardinerne fra i køkkenet
- Træk gardinerne fra i køkkenet

TS: Ruller gardinerne fra i køkkenet. Du er nu færdig med opgave 2.

Opgave 3

TS: Velkommen til dit stemmestyrede intelligente hjem. Hvad kan jeg hjælpe dig med?

FP:

- Sluk lamperne i soveværelset
- Sluk lyset i soveværelset

TS: Slukker lamperne i soveværelset.

FP:

- Indstil temperaturen til 22 grader.
- Sæt temperaturen til 22 grader.
- Ændre temperaturen til 22 grader.
- Temperaturen skal ændres til 22 grader.

TS: Indstiller temperaturen til 22 grader

FP:

- Afspil noget Jazzmusik
- Spil noget Jazzmusik
- Sæt noget Jazzmusik på

TS: Afspiller Jazzmusik

FP:

- Rul gardinerne for i stuen
- Træk gardinerne for i stuen

TS: Ruller gardinerne for i stuen. Du er nu færdig med træningsrunden.

A.10 Talespor til scenarie B

Scenarie B

Opgave 1

TS: Velkommen til din stemmestyrede bestilling af flyrejser. Kunne du tænke dig at bestille en rejse?

FP: Ja

TS: Hvad er din destination?

FP:

- Min destination er New York
- Jeg skal til New York
- Jeg vil til New York

TS: Sætter din destination til New York. Hvor ønsker du afgang fra?

FP:

- Jeg ønsker/vil have afgang fra Amsterdam.
- Jeg rejser fra Amsterdam
- Jeg tager afsted fra Amsterdam
-

TS: Sætter din afgang fra Amsterdam. Hvilken dato vil du rejse?

FP:

- Jeg rejser den 12. september/8 2016
- Jeg tager afsted den 12. september/8 2016
- Jeg vil afsted den 12. september/8 2016
- Jeg vil rejse den 12. september/8 2016
- Jeg skal afsted den 12. september/8 2016
- Jeg skal rejse den 12. september/8 2016
- Datoen jeg skal/vil rejse er den 12. september/8 2016

TS: Sætter datoen til 12. September 2016. Hvilket tidspunkt ønsker du afrejse?

FP:

- Jeg ønsker at afrejse kl. 06:45
- Jeg tager afsted kl. 06:45
- Jeg vil afsted kl. 06:45
- Jeg vil rejse den kl. 06:45
- Jeg skal afsted den kl. 06:45
- Jeg skal rejse den kl. 06:45
- Tidspunktet jeg vil rejse/ønsker afrejse er kl. 06:45

TS: Sætter tidspunkt for afrejsen til kl. 06:45. Hvor mange styk bagage ønsker du at medbringe?

FP:

- Jeg har/medbringer 3 styk bagage/3 tasker/3 kufferter
- Jeg vil gerne indtjekke 3 styk bagage/3 tasker/3 kufferter

TS: Sætter antal bagage til 3 styk. Din bestilling er nu lavet.

Opgave 2

TS: Velkommen til din stemmestyrede bestilling af flyrejser. Kunne du tænke dig at bestille en rejse?

FP: Ja

TS: Hvad er din destination?

FP:

- Min destination er Shanghai
- Jeg skal til Shanghai
- Jeg vil til Shanghai

TS: Sætter din destination til Shanghai. Hvor ønsker du afgang fra?

FP:

- Jeg ønsker/vil have afgang fra Frankfurt.
- Jeg rejser fra Frankfurt
- Jeg tager afsted fra Frankfurt

TS: Sætter din afgang fra Frankfurt. Hvilken dato vil du rejse?

FP:

- Jeg rejser den 25. september/8 2016
- Jeg tager afsted den 25. september/8 2016
- Jeg vil afsted den 25. september/8 2016
- Jeg vil rejse den 25. september/8 2016
- Jeg skal afsted den 25. september/8 2016
- Jeg skal rejse den 25. september/8 2016
- Datoen jeg skal/vil rejse er den 25. september/8 2016

TS: Sætter datoen til 25. September 2016. Hvilket tidspunkt ønsker du afrejse?

FP:

- Jeg ønsker at afrejse kl. 12:30
- Jeg tager afsted kl. 12:30
- Jeg vil afsted kl. 12:30

- Jeg vil rejse den kl. 12:30
- Jeg skal afsted den kl. 12:30
- Jeg skal rejse den kl. 12:30
- Tidspunktet jeg vil rejse/ønsker afrejse er kl. 12:30

TS: Sætter tidspunkt for afrejsen til kl. 12:30. Hvor mange styk bagage ønsker du at medbringe?

FP:

- Jeg har/medbringer 1 styk bagage/1 tasker/1 kufferter
- Jeg vil gerne indtjekke 1 styk bagage/1 tasker/1 kufferter

TS: Sætter antal bagage til 1 styk. Din bestilling er nu lavet.

Opgave 3

TS: Velkommen til din stemmestyrede bestilling af flyrejser. Kunne du tænke dig at bestille en rejse?

FP: Ja

TS: Hvad er din destination?

FP:

- Min destination er London
- Jeg skal til London
- Jeg vil til London

TS: Sætter din destination til London. Hvor ønsker du afgang fra?

FP:

- Jeg ønsker/vil have afgang fra Aalborg.
- Jeg rejser fra Aalborg
- Jeg tager afsted fra Aalborg

TS: Sætter din afgang fra Aalborg. Hvilken dato vil du rejse?

FP:

- Jeg rejser den 2. oktober/10 2016
- Jeg tager afsted den 2. oktober/10 2016
- Jeg vil afsted den 2. oktober/10 2016
- Jeg vil rejse den 2. oktober/10 2016
- Jeg skal afsted den 2. oktober/10 2016
- Jeg skal rejse den 2. oktober/10 2016
- Datoen jeg skal/vil rejse er den 2. oktober/10 2016

TS: Sætter datoen til 2. oktober 2016. Hvilket tidspunkt ønsker du afrejse?

FP:

- Jeg ønsker at afrejse kl. 17:00
- Jeg tager afsted kl. 17:00
- Jeg vil afsted kl. 17:00
- Jeg vil rejse kl. 17:00
- Jeg skal afsted kl. 17:00
- Jeg skal rejse kl. 17:00
- Tidspunktet jeg vil rejse/ønsker afrejse er kl. 17:00

TS: Sætter tidspunkt for afrejsen til kl. 17:00. Hvor mange styk bagage ønsker du at medbringe?

FP:

- Jeg har/medbringer 2 styk bagage/2 tasker/2 kufferter
- Jeg vil gerne indtjekke 2 styk bagage/2 tasker/2 kufferter

TS: Sætter antal bagage til 2 styk. Din bestilling er nu lavet.

A.11 Talespor til scenarie C

Scenarie C

Opgave 1

TS: Velkommen til din stemmestyrede kalender. Ønsker du at oprette en begivenhed?

FP: Ja

TS: Okay. Hvilken dato foregår begivenheden?

FP:

- Begivenheden foregår den 12. Oktober/10 2016
- Det er den 12. Oktober/10 2016
- Datoen er den 12. Oktober/10 2016

TS: Okay. Hvilket tidspunkt på dagen finder begivenheden sted?

FP:

- Begivenheden finder sted klokken 18:00
- Tidspunktet er klokken 18:00
- Det er klokken 18:00

TS: Okay. Hvad er emnet for begivenheden?

FP:

- Emnet er Fødselsdag
- Det er Fødselsdag
- Begivenheden er fødselsdag

TS: Okay. Hvad skal jeg påminde dig om til begivenheden?

FP:

- Påmind mig om at jeg skal købe en gave
- Husk mig på at jeg skal købe en gave
- Jeg skal købe en gave
- Lav en notifikation om at jeg skal købe en gave

TS: Okay. Din begivenhed er nu oprettet.

Opgave 2

TS: Velkommen til din stemmestyrede kalender. Ønsker du at oprette en begivenhed?

FP: Ja

TS: Okay. Hvilken dato foregår begivenheden?

FP:

- Begivenheden foregår den 7. December/12 2016
- Det er den 7. December/12 2016
- Datoen er den 7. December/12 2016

TS: Okay. Hvilket tidspunkt på dagen finder begivenheden sted?

FP:

- Begivenheden finder sted klokken 13:45
- Tidspunktet er klokken 13:45
- Det er klokken 13:45

TS: Okay. Hvad er emnet for begivenheden?

FP:

- Emnet er Bryllup
- Det er Bryllup
- Begivenheden er Bryllup

TS: Okay. Hvad skal jeg påminde dig om til begivenheden?

FP:

- Påmind mig om at jeg skal stryge tøj
- Husk mig på at jeg skal stryge tøj
- Jeg skal huske at stryge tøj
- Lav en notifikation om at jeg skal stryge tøj

TS: Okay. Din begivenhed er nu oprettet.

Opgave 3

TS: Velkommen til din stemmestyrede kalender. Ønsker du at oprette en begivenhed?

FP: Ja

TS: Okay. Hvilken dato foregår begivenheden?

FP:

- Begivenheden foregår den 27. Maj/5 2017
- Det er den 27. Maj/5 2017
- Datoen er den 27. Maj/5 2017

TS: Okay. Hvilket tidspunkt på dagen finder begivenheden sted?

FP:

- Begivenheden finder sted klokken 10:00

- Tidspunktet er klokken 10:00
- Det er klokken 10:00

TS: Okay. Hvad er emnet for begivenheden?

FP:

- Emnet er Karneval
- Det er Karneval
- Begivenheden er Karneval

TS: Okay. Hvad skal jeg påminde dig om til begivenheden?

FP:

- Påmind mig om at jeg skal finde et kostume
- Husk mig på at jeg skal finde et kostume
- Jeg skal huske at finde et kostume
- Lav en notifikation om at jeg skal finde et kostume

TS: Okay. Din begivenhed er nu oprettet.