

Del 5 – Brukertestning av Multimodal UI pilot

Prosjektnummer: 183215/S10

Prosjekttittel: SMUDI

Skrevet av: Miriam Begnum, Cathy Kalvenes, Bjørn Guldbrandsen

Sist oppdatert: 02.12.2011

Innholdsfortegnelse

1. Bakgrunn	3
1.1. Utarbeidelse av piloten	3
1.1.1 Definerings av taledialog	3
1.2. Testplan	5
1.3. Rekruttering av testpersoner	5
1.4. Gjennomføring av brukertester.....	5
1.4.1 Observasjon.....	6
1.4.2 Intervju	7
1.5. Analyse	7
1.6. Resultater	7
1.6.1 Synshemmede	8
1.6.2 Bevegelseshemmede.....	12
1.6.3 Dyslektikere og personer med lese- og skrivevansker	15
1.7. Oppsummering av hovedfunn	18
1.8. Konklusjon	19
Vedlegg A: Nettbasert multimodalt dialogsystem for værvarslingsinformasjon.....	20
Dialogsystemets funksjonalitet	21
Direkte dialog	21
Kort oversiktsvarsel	23
Time for time.....	23
Helgevarsel	24
Langtidsvarsel.....	24
Vedlegg B: Teknisk spesifikasjon og dialogeksempel	25
Vedlegg C: Informasjons- og samtykkeerklæring.....	26
Bakgrunn	26
Praktisk informasjon for deltakere.....	26
Prosjektet	26
Personvern	27
Samtykke	27
Vedlegg D: Observasjonsskjema og intervjuguide	28
Intervju	28

1. Bakgrunn

Samfunnsutviklingen går raskt i retning av at alle må "betjene seg selv", og det er derfor viktig å finne frem til dialogsystemer som kan brukes av alle. I forprosjektet Stemint (<http://www.medialt.no/stemint/85.aspx>) kom det fram at brukere med dysleksi, bevegelseshemming og synshemming ønsket å kunne bruke talestyring i dialoger på nett. Det finnes flere talebaserte brukergrensesnitt (f.eks. telefonavisa), men kartleggingen i Stemint-prosjektet tyder på at brukerne ønsker seg "rikere" brukergrensesnitt. Blant annet ønsket de å kunne benytte talebaserte dialoger i utfyllinger av skjemaer og til å hente ut informasjon fra nettet. De ulike brukergruppens behov var sammenfallende – noe som tyder på at dette kan være hensiktsmessige grensesnitt for helt generelt bruk. Hvordan brukerne opplever en multimodal dialog, hvilke modaliteter som velges og foretrekkes av brukere som kan benytte flere modaliteter og hvorvidt man kan lage systemer der alle kombinasjoner av modaliteter er tilstrekkelig (mus/tale, tale/tale, punkt/tale, ...) er eksempler på forskningsspørsmål det var interessant å se nærmere på.

En pilot med talebasert dialog i tillegg til andre modaliteter ble derfor utviklet. Værdata fra yr ble benyttet i en nettbasert værøktjeneste; der tale var lagt til som input og output modaliteter i et forenklet visuelt brukergrensesnitt. Denne multimodale piloten ble deretter brukertestet. Utviklingen av den multimodale piloten og testingen av denne ble gjennomført som en del av forskningsprosjektet Stemstyring i Multimodal Dialog (www.smudi.no). Denne rapporten beskriver utarbeidelse, testing og analyse av piloten.

1.1. Utarbeidelse av piloten

Den svenske samarbeidspartneren Veridict har stått for teknisk utvikling av piloten. Veridict er utvikler av talegjenkjenningmotor på norsk i SMUDI prosjektet.

MediaLT hadde ansvar for spesifisering av piloten. Brukergrensesnittet ble utviklet hovedsakelig med tanke på å legge til rette for en god taleinteraksjon, mens det grafiske grensesnittet og tastaturtilgjengelighet var sekundært. I spesifiseringsarbeidet ble utarbeidelsen av taledialog prioritert fremfor visuelt grensesnitt og tastaturtilgjengelighet. Årsaken var at taleinteraksjon var fokus i piloten, og den minst kjente interaksjonsmetoden.

1.1.1 Definerings av taledialog

Målet med dialogen var å gi brukeren frihet i interaksjonen, og initiativ ble derfor lagt hos brukeren fremfor hos systemet. I stedet for å spesifisere opp en fastlagt dialog, ble det derfor bestemt at systemet skulle lytte etter, og gjenkjenne, et utvalg ordfraser (nøkkelord) i en strøm av lyder.

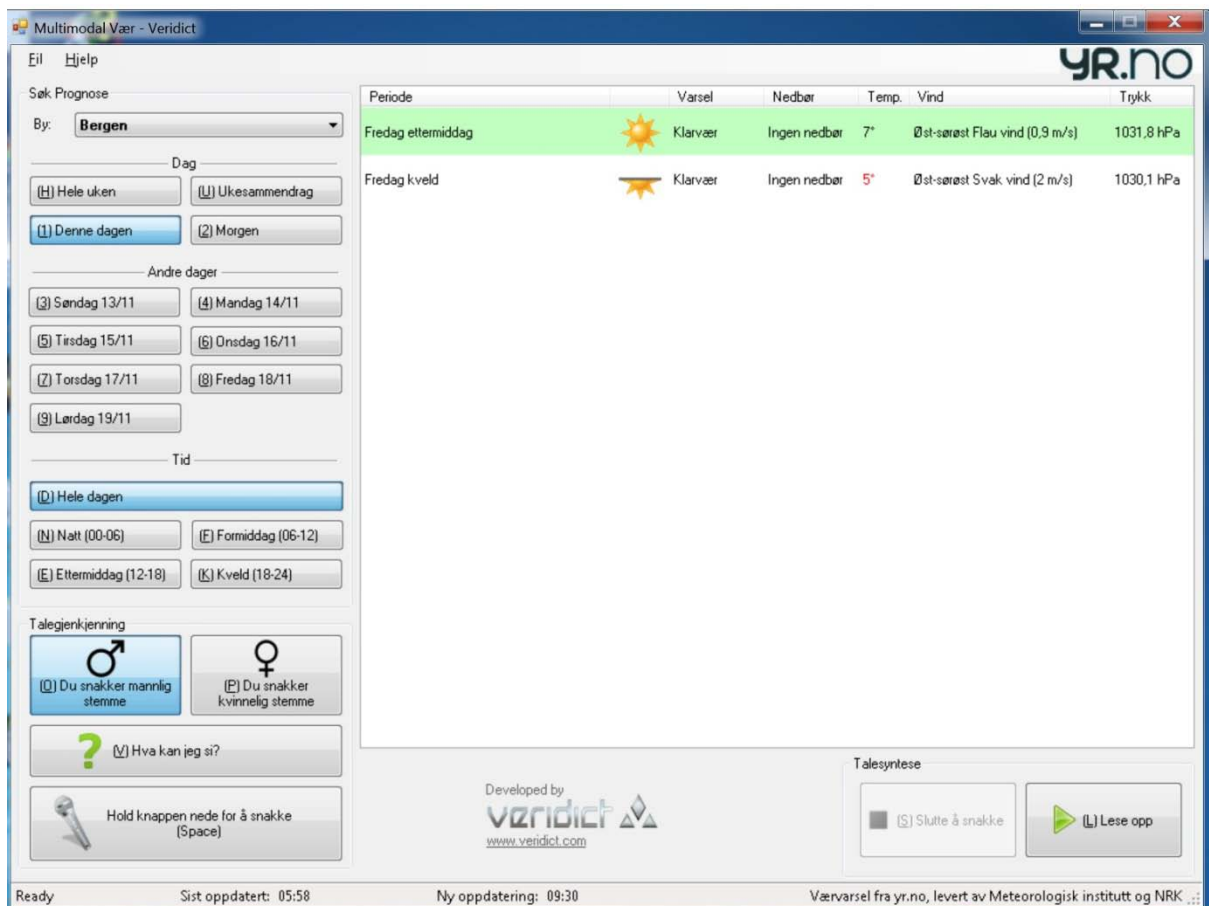
Første spesifisering av dialogsystemet finnes som Vedlegg A, og ble godkjent av Veridict som gjennomførbar i forkant av utviklingsprosessen. Plattformen til piloten er Microsoft Windows 7 - 64 bit. Vedlegg B er eksemplifisering av dialog samt tekniske spesifikasjoner. Selv om dialogen av utviklingsmessige årsaker ikke ble implementert nøyaktig etter spesifikasjonene, ble de vesentlige aspektene ved dialogen beholdt; brukerfrihet og nøkkelord. Hensikten med piloten var ikke knyttet

opp til en spesifikk taledialog, og det ble derfor ikke lagt stor vekt på å unngå endringer i dialogdetaljer.

Endringene i endelig pilot besto hovedsakelig i at den introduserende taledialogen ble fjernet, og dialogen ble noe forenklet, med bl.a. begrenset feedback av bekreftelse på valg fra talesyntesen. Dette gjør at piloten ikke er optimal å benytte for sterkt synshemmede og blinde uten samtidig bruk av skjermleserfunksjonalitet. Hjelpfunksjonaliteten "Hva kan jeg si" ble imidlertid beholdt for å veie opp for noe av dette, slik at mulighetene i grensesnittet vil være synlige for alle brukere.

Earcone-ikoner med lyder for ulike væertyper ble ikke inkludert av ressursmessige årsaker. Det endelige grensesnittet viser visuelt resultatene av et værsøk i tabellform ved noe grafikk, enkel tekst og enkle tall, med en rullgardinmeny for valg av sted, og ellers knapper for valg av tidspunkt og dag.

Figur 1: Endelig visuelt grensesnitt



Den endelige multimodale piloten ble interntestet av seende, svaksynte og blinde personer, både med hjelpemiddelteknologiene JAWS og ZoomText og uten bruk av hjelpemidler, for å sikre at løsningen var kompatibel med relevant hjelpemiddelteknologi, forståelig og anvendelig.

1.2. Testplan

En testplan ble utformet med den hensikt å undersøke hvordan talestyring i dialog fungerer for målgruppen; frittstående eller i kombinasjon med andre måter å styre applikasjonen på. Målgruppen for testingen var personer med dysleksi, personer som har problemer med å bruke mus og/eller tastatur og synshemmede personer.

Det overordnede spørsmålet var om talebaserte dialoger i grensesnitt kunne være hensiktsmessig for å bedre brukbarheten for funksjonshemmede brukergrupper; jamfør SMUDI prosjektets målsetting: "Utvikle multimodale grensesnitt med nødvendig redundans for å sikre universell tilgjengelighet til styring av klientutstyr og nettbaserte dialogsystemer."

Hensikten med testingen av piloten ble imidlertid mer spesifikt definert som å undersøke hvilke interaksjonsmetode bevegelseshemmede, dyslektikere og synshemmede foretrekker og har nytte av i applikasjonens multimodale brukergrensesnitt. Dette er operasjonalisert gjennom følgende forskningsspørsmål:

1. **Hvilke interaksjonsmuligheter** benyttes til input i de tre brukergruppene, og **hvordan oppleves** disse?
2. **Hvilke interaksjonsmuligheter** benyttes til output i de tre brukergruppene, og **hvordan oppleves** disse?
3. **Hvor raskt** utføres de forskjellige oppgavene i de tre brukergruppene, og hvordan påvirkes hurtigheten av type interaksjonsmuligheter som velges?
4. Hva er **totalopplevelsen** av det multimodale brukergrensesnittet i de tre brukergruppene, og hvordan ser testerne på nytteverdien av slike grensesnitt?

1.3. Rekruttering av testpersoner

Målsetningen var å få rundt 10 testpersoner i hver kategori. Testene ble rundet av etter testing med 10 dyslektikere, 8 bevegelseshemmede og 9 synshemmede. Rekrutteringen skjedde i stor grad via kontaktnettverk, og det ble rekruttert en stor bredde av ulike brukere med hensyn på type funksjonshemming, grad av funksjonshemming, kjennskap til PC, kjennskap til hjelpemiddelteknologi, type hjelpemiddelteknologi, alder, kjønn, bosted og yrkesstatus.

Som incitament til deltakelse ble fem deltakere trukket ut som vinnere av gavekort til en verdi av 1000 NOK.

1.4. Gjennomføring av brukertester

To testledere sto for gjennomføring av hver brukertest. Den ene av disse – testveileder – var den samme gjennom alle brukertestene. Den andre – ansvarlig for dokumentasjon – varierte mellom to personer.

Hver brukertester mottok først en gjennomgang av informasjons- og samtykkeerklæringen (Vedlegg C i rapporten), og denne erklæringen ble undertegnet før testing startet. Deretter fikk testpersonen en rask innføring i hvordan applikasjonen brukes, på i underkant av 10 minutter. Etter innføringen ble testpersonen gitt 8 forskjellige oppgaver å utføre. Under gjennomføringen av testen observerte testlederne hvor raskt testpersonen utførte de forskjellige oppgavene og hvilke interaksjonsmuligheter som ble benyttet, til input og output. I tillegg til observasjon av gjennomføringen, ble korte semi-strukturerte intervjuer holdt i etterkant av testingen. Hver test tok ca. 1 time.

1.4.1 Observasjon

Et observasjonsskjema ble benyttet, der bakgrunnsdata om testpersonene ble registrert, slik som navn, alder, hjelpemidler og problematikk i forhold til PC-bruk. Tiden benyttet per oppgave, hva slags input og output metode som ble benyttet i hver oppgave og annen relevant informasjon ble også notert i observasjonsskjemaet. Observasjonsskjemaet finnes som Vedlegg D, i tillegg til intervjuguide.

Opgavene ble arrangert på en slik måte at testpersoner måtte benytte alle tilgjengelige modaliteter som input for å fullføre oppgaver. Input-modalitet var i de to første oppgavene valgfri, slik at man kunne observere hvilken løsning testpersonen var mest komfortabel med umiddelbart etter innføring i begge metodene. I de neste fire oppgavene var modaliteten spesifisert, hhv. mus/tastatur og tale, for å tillate observasjon av disse interaksjonsmulighetene. De siste to oppgavene var igjen valgfri, og indikerer hvilken input metode som foretrekkes etter at testerne har benyttet alle tilgjengelige interaksjonsmetoder. De ulike outputmetodene (skjerm/tale/leselist) var valgfrie i alle oppgavene.

De første syv oppgavene var nærmest identiske, og besto av enkle værøksøk på en gitt dag, på et gitt tidspunkt i en gitt by. Den siste oppgaven var noe mer komplisert, med spørsmål om forskjell i temperatur mellom to værøksøk, for å sikre at brukerne mestret grensesnittet og modalitetene som informasjonsbærende metoder. I tillegg kunne man da få et inntrykk av foretrukne modalitetsvalg når testerne ble presentert med et mer komplekst søk og med større mulighet for eventuelle feil.

Opgavene var som følger:

1. Hvordan blir været lørdag kveld i Oslo?
2. Hvordan blir været mandag ettermiddag i Bergen?
3. Bruk tastatur eller mus til å finne været tirsdag kveld i Trondheim.
(Oppgaven utføres kun dersom testerne kan benytte mus eller tastatur.)
4. Bruk tastatur eller mus til å finne ut hvordan været blir søndag formiddag i Stavanger.
(Oppgaven utføres kun dersom testerne kan benytte mus eller tastatur.)
5. Bruk stemmestyring til å finne været mandag natt i Fredrikstad.
6. Bruk stemmestyring til å finne været søndag ettermiddag i Tromsø.
7. Hvordan blir været i torsdag kveld Drammen?
8. Vis været for hele dagen på lørdag i Fredrikstad og finn forskjellen i temperatur mellom natt og kveld.

1.4.2 Intervju

Intervjuene søkte å måle brukernes opplevelse av det multimodale grensesnittet, og hvilke interaksjonsmuligheter som ble foretrukket. Det ble benyttet en semi-strukturert intervjuguide. Intervjuspørsmålene var som følger:

1. Hvilken input metode og output metode opplevde du som mest effektiv og mest behagelig å bruke?
2. Synes du talegjenkjenningen fungerte tilfredsstillende?
3. Var det noe du syntes var spesielt bra eller spesielt dårlig med systemet?
4. Vil slike systemer gjøre det enklere for deg å få tilgang til informasjon fra internett? Begrunn svaret.
5. Kan du si noe om når du foretrekker å bruke ulike måter å gi input og få output på?

Oppgave 5 ble ved behov utdypet med eksempler, som; at noen kan foretrekke å lese telefonnummer med leselist, men å få lange tekster opplest med tale; at noen vil se grafiske symboler for været, men få tekst lest opp; og at noen vil gi kommandoer med tale, men selv se resultater på skjerm.

1.5. Analyse

Som tidligere skissert, skulle testingen undersøke hvilke modaliteter de tre brukergruppene foretrakk i et multimodalt grensesnitt som inkluderer tale, deres tidsbruk ved bruk av modalitetene til værøøk og hvordan bruken av modalitetene og piloten som helhet ble opplevd. Resultatene ble analysert med hensyn på dette.

Testene og analysen er tilstrebet utformet nøytralt i forhold til antakelser om hvorvidt tale ville foretrekkes som input. Det kan imidlertid påvirke resultatene noe at piloten ikke er optimalt tilpasset bruk med kun tastatur som inputmetode. Muligheter for hurtigtaster er ikke inkludert i testene, selv om disse foreligger. Dette antas å kunne ha påvirket opplevelsen av tastatur som inputmetode noe. Det var heller ingen hypoteser knyttet til hvilken output brukerne ville foretrekke.

1.6. Resultater

I piloten krevde en systemfeil at værøøket måtte stilles på en bestemt måte; talegjenkjenningen ble svekket når stedsnavn ikke ble sagt til slutt i en talekommando. Dette slet flere i brukergruppen med å mestre. Det var vanskelig for dem å huske den "riktig" rekkefølgen, muligens delvis fordi stedsvalg visuelt ligger først på siden. Dette førte til noen problemer, og ble kommentert som negativt. Noen testere pekte også på manglende muligheter for forstørrelse av visuell informasjon i piloten.

Gjennomsnittlig tid benyttet på værøøk er kun kalkulert fra oppgave 1-7. Oppgave 8 er ikke inkludert i tabellen, da denne oppgaven er ulik de andre, og mer avansert. Oppgave 8 er inkludert i testplanen for å observere bruk og hensiktsmessighet av modaliteter til mer kompleks informasjonshenting. Tidsbruken her kan derfor vanskelig sammenlignes. Det må tas høyde for metode for informasjonshenting av tidsopplysninger når snittidene vurderes; stoppeklokke ble startet og

stoppet manuelt under testobservasjonene. Små variasjoner i start og stopptidspunkter må derfor påberegnes. Dette må antas å ha gitt dataene en usikkerhetsfaktor på opptil et par sekunder.

1.6.1 Synshemmede

Denne testgruppen omfattet tre svaksynte, to sterkt svaksynte og fire blinde i alderen 24-60 år.

De fire blinde testerne var skjermleserbrukere (testpersoner 2, 3, 6 og 9). Kompetansen varierte imidlertid veldig. Testperson 6 både leste punkt på leselist og hadde mye erfaring med bruk av PC og skjermleser. Testperson 3 hadde en del års erfaring med å bruke PC med skjermleser, men leste ikke punkt, mens testperson 9 var en rask punktleser som foretrakk å bruke leselist fremfor tale. Testperson 2 hadde lav datakompetanse, og kunne ikke tastaturlayout.

De to sterk svaksynte (testpersoner 4 og 5) hadde også lite erfaring med bruk av PC, og kunne heller ikke layout på tastaturet. Begge var til utredning for hjelpemidler da testene ble gjennomført. Testperson 4 brukte skjermforstørrelser med tale.

Av de tre moderat svaksynte var også en person under utredning (testperson 8). Testere 8 og 7 var vant med å bruke mus og hadde i liten grad benyttet tastatur for å styre PC. Den siste, testperson 1, benyttet imidlertid både skjermforstørrelser med tale, tastatur og mus til å styre PC-en.

Tabell 1: Oversikt over testpersoner med synshemminger

	Alder	Beskrivelse av funksjonsproblematikk	Hjelpemidler i forhold til PC-bruk
Testperson 1	24	Svaksynt	Skjermforstørrelser/skjermleser (Supernova) og VoiceOver.
Testperson 2	32	Blind	Skjermleser (Jaws).
Testperson 3	49	Blind	Skjermleser (Jaws).
Testperson 4	48	Sterkt svaksynt (RP)	Skjermforstørrelser (ZoomText).
Testperson 5	48	Sterkt svaksynt	Til utredning.
Testperson 6	51	Blind	Skjermleser med leselist (Jaws).
Testperson 7	44	Redusert synsfelt (synsfeltutfall)	Ingen hjelpemidler
Testperson 8	60	Svaksynt	Til utredning
Testperson 9	29	Blind	Skjermleser med leselist (Jaws).

1.6.1.1 Valg av input

Tabell 2 viser antall testere per benyttede input metode. På de valgfrie oppgavene (oppgave 1, 2, 7 og 8) ble tale så godt som alltid foretrukket fremfor tastatur som input metode, mens det for mus versus tale er mer uklart hva som foretrekkes.

Tabell 2: Antall testere per benyttede inputmetode (antall/antall med mulighet)

	Oppgave 1	Oppgave 2	Oppgave 3	Oppgave 4	Oppgave 5	Oppgave 6	Oppgave 7	Oppgave 8
Tale	7/9	7/9	-	-	9/9	9/9	7/9	6/9
Tale + tastatur	0/4	1/4	-	-	-	-	0/4	0/4

Tastatur	0/4	0/4	4/4	4/4	-	-	0/4	1/4
Mus	1/3	1/3	2/3	2/3	-	-	2/3	2/3
Mus + tastatur	1/1	0/1	0/1	0/1	-	-	0/1	0/1

De svaksynte testpersonene 1, 7 og 8 hadde mulighet for å bruke datamus som input metode. Sterkt synshemmede personer er imidlertid uten mulighet for å benytte mus. Tastatur inn og leselist ut var alternative interaksjonsmodaliteter til tale for de seks blinde og sterkt svaksynte testpersonene, hvorav tre ikke hadde kompetanse nok til å benytte tastatur som input metode. Tre personer kunne dermed hverken benytte mus eller tastatur, og kunne ikke utføre oppgavene 3 og 4.

Det var dermed i testgruppen tre personer som *kun* kunne benytte tale som input, tre personer som kunne benytte tastatur og tale som input, to personer som kunne benytte mus og tale som input og en som kunne benytte både tastatur, mus og tale.

1.6.1.2 Valg av output

Tabell 3 oppsummerer bruken av output metoder.

Tre svaksynte hadde nok syn til å få nytte av visuelt output på skjerm, hvorav to av dem (testere 1 og 7) for det meste valgte skjerm i kombinasjon med taleoutput. Til to oppgaver valgte en av disse kun å benytte tale.

Tabell 3: Antall testere per benyttede outputmetode (antall/antall med mulighet)

	Oppgave 1	Oppgave 2	Oppgave 3	Oppgave 4	Oppgave 5	Oppgave 6	Oppgave 7	Oppgave 8
Kun tale	6/9	5/9	4/9	4/9	7/9	6/9	7/9	6/9
Tale + leselist	0/2	1/2	2/2	2/2	0/2	0/2	0/2	0/2
Tale + skjerm	2/3	2/3	2/3	2/3	1/3	2/3	1/3	2/3
Kun skjerm	1/3	1/3	1/3	1/3	1/3	1/3	1/3	1/3

De øvrige seks testpersonene benyttet taleoutput fra applikasjonen under testen. Fire av disse benyttet kun tale fra piloten, mens to kombinerte med tale fra skjermleser (testere 3 og 6). De to som kunne lese punkt, benyttet sporadisk leselist som støtte. Bruken av leselist som output metode hang sammen med hvilken input metode som ble benyttet – der tastatur var blitt benyttet og hendene lå plassert ved leselisten, ble listen gjerne brukt som støtte til tale.

“Output der foretrekker jeg list fordi det er det jeg er vant til. Men hvis fokuset på lista ikke er der informasjonen er, foretrekker jeg tale. Så det kommer an på hva jeg bruker som input.”

– Testperson 9

1.6.1.3 Opplevelse av tale som input

“Veldig lettvinnt med talestyring!” – Testperson 6

Generelt mente alle at talegjenkjenningen fungerte tilfredsstillende, og spesielt at det var enkelt å komme direkte til ønsket informasjon. Samtlige svaksynte vant med hjelpemiddelteknologi og talesyntese valgte tale som input i de valgfrie oppgavene, i alt syv testere. Alle syv uttrykte at talestyring var mer effektiv og behagelig å bruke enn tastatur/mus.

Tre av testerne hadde svært lav datakompetanse. Det er interessant at alle disse var svært fornøyde med tale input, og beskrev det som både lett å lære og å bruke. To var fremmedspråklige og uttalelse av talekommandoene var ikke helt perfekt, men de opplevde allikevel at talegjenkjenningen fungerte tilfredsstillende.

Testpersonene 7 og 8 brukte ikke talestyring i de valgfrie oppgavene, og ga uttrykk for at taleinput er nytt og uvant. For disse to var datamus den naturlige input metoden, og et visuelt grensesnitt med knapper fungerte godt for dem.

1.6.1.4 Opplevelse av tale som output

“... raskt og effektiv å få svar.” – Testperson 3

Alle testerne unntatt en valgte å benytte tale output – enten alene eller i kombinasjon med mus/leselist. Det virket naturlig å bruke tale output modaliteten når talekommandoer var benyttet til input.

Tre testpersoner hadde lang erfaring i å bruke talesynteser i arbeid, og disse var de aller mest positive til tale som output.

De to som benyttet tale fra applikasjonen sammen med skjermleser nevnte det var irriterende når både skjermleseren og piloten snakker samtidig.

1.6.1.5 Nytt av talestyrte dialogsystemer

Blinde og sterk svaksynte, inkludert nye databrukere, ga tydelig uttrykk for at talestyring ville gjøre det mye enklere for dem å hente ulik informasjon fra Internett.

“... du kan unngå mye heft. Du unngår grafikk. På nyhetssider kan du komme rett til artikkelen og slippe å høre søkefelt og andre ting først.” – Testperson 3

1.6.1.6 Tid brukt

Seks av ni testpersoner ønsket å benytte kun talemodaliteter til værøøk. Disse var alle blant de testpersonene med sterkest synshemming.

De tre øvrige testerne ønsket skjerm som output metode – enten sammen med tale eller alene. To benyttet mus som input metode – sammen med tale eller alene. Kun en person hadde ingen

talemodaliteter i sin prefererte interaksjonsform; testperson 7. Hun preferer mus og tastatur, og både tidsbruk og modaliteter sammenfaller med seende testere.

I søk der skjerm var primær output metode, virket enkle søk å ta under 10 sekunder uavhengig av om input metode var mus eller tale. Der tale var en viktig output modalitet, og testpersonene ventet på opplesingen, tok søkene derimot rundt 20 sekunder (her 19,12 sekunder i snitt, med gjennomsnittlig tid på enkelt informasjonssøk ved bruk av kun talemodaliteter på 22,5 sekunder).

Når testerne erstattet talekommandoer/mus med tastatur som input metode, ble søkene forlenget med ytterligere ca. 10 sekunder til et snitt rundt 30 sekunder (her 32,75 sekunder). Mot tastatur som input metode sparer tale input testerne mellom 7-17 sekunder, og i snitt 11,6 sekunder.

I tabell 4 finner du gjennomsnittlig tidsbruk på enkle værøsk og foretrukne interaksjonsmetoder.

Tabell 4: Tidsbruk per testperson, relatert til interaksjonsmetoder

	Foretrukket metode	Gjennomsnittlig tid (sekunder)	Øvrige metoder	Gjennomsnittlig tid (sekunder)
Testperson 1	Tale inn, Skjerm + tale ut	18,7	Tastatur inn, Skjerm + tale ut	36
Testperson 2	Tale inn, Tale ut	20	N/A	N/A
Testperson 3	Tale inn, Tale ut	21,2	Tastatur inn, Tale ut	31,5
Testperson 4	Tale inn, Tale ut	29,5	N/A	N/A
Testperson 5	Tale inn, Tale ut	21	N/A	N/A
Testperson 6	Tale inn, Tale ut	20,7	Tastatur inn, Tale + leselist ut	28
Testperson 7	Mus inn, Skjerm ut	9,4	Tale inn, Skjerm ut	7
Testperson 8	Mus inn, Skjerm + tale ut	17,5	Tale inn, Skjerm + tale ut	24,7
Testperson 9	Tale inn, Tale ut	19,8	Tastatur inn, Tale + leselist ut	35,5

Oppgave 1 fra testperson 6, oppgave 2 fra testperson 1 og oppgave 5 fra testperson 5 ble utelatte fra utregningen på grunn av talegjenkjenningsfeil som skyldtes piloten.

Oppgave 1 fra testperson 4 ble utelatt da hun måtte ha hjelp for å benytte talemodaliteten – hun var noe ustø i bruken av tale i de to første oppgavene og måtte bl.a. få svaret lest opp flere ganger. Misforståelser i forhold til bruk av talemodaliteten og kommunikasjon med testleder gjorde også at oppgavene 7 fra tester 2 og 1 og 2 fra tester 9 utelates.

1.6.1.7 Totalopplevelse

Alle testerne unntatt en foretrakk å benytte minst en talemodalitet i interaksjon med piloten.

Mus prefereres tilsynelatende over tale input, mens tale foretrekkes fremfor tastatur. Enkle informasjonssøk effektiviseres ved tale som input, og effektiviseres nokså vesentlig i forhold til tastatur som input metode.

Når det gjelder output benyttes gjerne alle tilgjengelige modaliteter sammen; tale som tillegg til leselist eller skjerm. Dette samstemmer med informasjonen fra STEMINT fokusgruppene, der synshemmede uttrykte ønske om å få spørsmål og svare på disse, men i tillegg bruke en leselist for å kontrollere, lese stedsnavn og prisinfo osv.

De personene som ikke har muligheter til å benytte standard modaliteter som tastatur, mus og skjerm, oppnår like rask interaksjon ved bruk av talemodaliteter som mer erfarne og/eller mindre synshemmede medtestere – selv ved førstegangs bruk. Det tar imidlertid lenger tid å lytte til en taleremse enn å få informasjon visuelt.

Talemodalitetene gir altså en svært nyttig redundans i grensesnittet for synshemmede brukere, og viser seg som et svært nyttig verktøy både som selvstendige modaliteter og som tillegg.

1.6.2 Bevegelseshemmede

De åtte bevegelseshemmede testpersonene var i alderen 30 til 67 år. Det var en svært heterogen gruppe med hensyn til alvorlighetsgrad og type funksjonshemming; spesifikt knyttet til redusert arm/håndfunksjon som hindrer bruk av standard input løsninger (mus eller tastatur). Tabell 5 viser en oversikt over testpersonene.

Tabell 5: Oversikt over testpersoner med bevegelseshemminger

	Alder	Beskrivelse av funksjonsproblematikk	Hjelpemidler i forhold til PC-bruk
Testperson 11	54	Multipel Sklerose; kan ikke bevege hender/fingre. Har klar tale.	VOMOTE talestyring av PC. Kan ikke benytte bryter eller tastatur.
Testperson 12	30	AMC (Arthrogryposis multiplex congenita); medfødt bøyde og stive ledd, feilstillinger i armer og ben med armene rotert og begrensning i bevegelse.	Munnpinne på tastatur, kulemus som styres med munn/kinn/hake.
Testperson 13	67	Polioskade; sterkt redusert armfunksjon.	Fottastatur og fotbetjent mus. I testen: pekeplate og to fingre på standard tastatur.
Testperson 14	51	Muskelsyke; svekket muskulatur.	Skjermtastatur, Textpilot.
Testperson 15	43	Gillain Barres-syndrom; mangler leddfangst (tilbakemelding fra nervesystemet om hva hun gjør), ukontrollerte bevegelser.	Stort tastatur og kulemus.
Testperson 16	60	Muskelsyke; lammelser i hender/føtter.	(Mikrolett) bryter, hodemus.
Testperson 17	49	Medfødt redusert motorisk funksjon i høyre side; overbelastning i venstre arm, svekket syn.	Større skjerm (21 ”), kulemus.
Testperson 18	55	Nedsatt bevegelse, sittehemming	PC på stativ, Dragon, kulemus.

Seks av de åtte testerne ville normalt brukt museløsninger som interaksjonsmetode i tilsvarende grensesnitt, mens de øvrige to benyttet talestyringssystemer. Alle testpersonene hadde mulighet for å benytte standard output på skjerm med unntak av testperson 17 som hadde redusert syn.

Testene ble ikke spesialtilpasset for personer avhengige av avansert hjelpemiddelteknologi som erstatning for tradisjonelt bruk av tastatur og mus, for eksempel skanning og hodemus. En forbedring av testene med tanke på input, ville derfor vært å installere piloten på brukernes egne PC-er; slik at testerne hadde hatt tilgang til tale input/output i tillegg til de hjelpemidlene de allerede benyttet.

1.6.2.1 Valg av input

Alle de åtte bevegelseshemmede testerne valgte å benytte tale som input i alle de fire "frie" oppgavene.

Kun tre av de åtte hadde muligheter for å benytte mus eller tastatur; testere 13 og 14 hadde tilstrekkelig arm/håndfunksjon til å bruke standard museløsninger (touchpad eller perifer mus) i testene, og testperson 12 hadde tilgang til alternativ museløsning (munnpinne samt kulemus som kunne styres med munn/kinn/hake) på egen PC i testene. Alle disse benyttet mus som input.

De øvrige 5 testerne hadde ikke mulighet til å teste piloten med annet enn tale; fire hadde ikke tilgang til alternative løsninger under testen og ulike typer muskelsykdommer som forhindret dem fra å utføre oppgaver med standard mus og tastatur, og den femte testpersonen hadde glemt lesebrillene hjemme og var forhindret fra å benytte mus/tastatur grunnet redusert syn. For disse fem brukerne ble oppgavene 3 og 4 sett bort fra.

1.6.2.2 Valg av output

Alle de syv med mulighet for å se informasjon på skjerm benyttet dette som output. To personer hadde ikke mulighet for tale ut fra piloten.

1.6.2.3 Opplevelse av tale som input

"Du bare sier det du vil ha greie på." – Testperson 16

Syv av testerne syntes at talegjenkjenningen fungerer tilfredsstillende. Flere var imponert over løsningen, og spesielt måten å interagere med et system med taleinput på. Alle testerne hadde en intuitiv forståelse av hvordan taleinput skulle brukes; hvordan de skulle snakke med systemet, og opplevde at dialogen med piloten følte naturlig.

Tale ble beskrevet som mindre anstrengende å benytte enn alternative interaksjonsmetoder, både fysisk og når det kommer til generell brukervennlighet; "enklere vei til muligheter i grensesnittet" og "mindre plundrete". For eksempel forklarer testperson 15 at siden hun har problemer med "finstyring av hender og fingre", er tale et meget godt alternativ til tastatur og mus.

To testpersoner hadde imidlertid problemer med å rette opp feil når talegjenkjenningen ikke fungerte slik de forventet, og en av disse uttrykte skepsis ovenfor løsningen.

1.6.2.4 Opplevelse av tale som output

”Er det ’høytaler-mulighet’ velger jeg det – på Word dokumenter for eksempel. For da kan jeg legge meg litt tilbake (han senker ryggen på rullestolen), så gjør teknologien det jeg egentlig skulle gjøre. Så kan jeg bare ta imot.” – Testperson 17

Tale som output på resultatet av vørsøket ble lite brukt i denne gruppen. Flere mente imidlertid det hadde vært fint med mer auditiv tilbakemelding på det søket systemet oppfattet ved tale som input.

Testperson 17 benyttet kun tale som output, da han ikke kunne nyttiggjøre seg av den informasjonen på skjerm. Han har svekket syn og liker i likhet med testperson 18 å veksle mellom visuell/auditiv input – både lese på skjerm og lytte samtidig. Hans vurdering av tale som output var at:

- For mye informasjon gis
- Informasjonen gis for tett; i et for raskt tempo
- Det mangler naturlige pauser i taleflyten
- Informasjonen gis i uhensiktsmessig rekkefølge

På basis av dette foreslo han muligheter i grensesnittet for å velge vekk informasjon, senke lesehastigheten eller gi muligheter for å justere denne, benytte en stemme med mer naturlig betoning og pause eller gi muligheter for å velge blant ulike stemmer (for eksempel kvinne/mann) og vurdere rekkefølgen på informasjonen som skal leses slik at viktig informasjon kommer tidlig, at ikke to tall leses nær hverandre osv.

1.6.2.5 Nyttien av talestyrte dialogsystemer

”Man må ha ett eller annet redskap for å betjene en PC, og tale er et glimrende slikt redskap!” – Testperson 15

På spørsmål om systemer med tale vil gjøre det enklere å få tilgang til informasjon fra Internett, svarer flere testere bekreftende på dette. Eksempler på nyttige løsninger nevnt er;

- Situasjoner der spesialutstyr ikke er enkelt tilgjengelig, for eksempel på reiser, ville talestyring av grensesnitt være nyttig.
- Applikasjoner som snarveier på skrivebordet; benytte disse i stedet for nettsidene de tilsvare.
- Å kjøre slike applikasjoner i bakgrunnen med tale output uten visuelt fokus, for eksempel for å få værvarsel mens man gjør lager arbeidsplan for assistenten i Word og planlegger ukeaktiviteter.
- Talegrensesnitt i e-postsystemer og ved dokumentbehandling.

Tester 17 fokuserer på at mulighetene for tale i slike grensesnitt må være allment tilgjengelige – dvs. ikke kreve ekstra utstyr. Han mener enklere interaksjonsformer vil være nyttig for alle; ”vi liker det (alle) behagelig”.

1.6.2.6 Tid brukt

I tabell 6 finnes gjennomsnittlig tidsbruk på enkle værøsk. Gjennomsnittlig tid benyttet på enkle værøsk med mus input og skjerm output er 14,5 sekunder. Gjennomsnittlig tid benyttet på enkle værøsk med tale input og skjerm output er 8,7 sekunder – altså 5,8 sekunder raskere enn ved mus.

Tabell 6: Tidsbruk per testperson, relatert til interaksjonsmetoder

	Foretrukkede interaksjonsmetoder	Gjennomsnittlig tidsbruk med disse (sekunder)	Øvrige interaksjonsmetoder	Gjennomsnittlig tidsbruk med disse (sekunder)
Testperson 11	Tale inn, Skjerm ut	9,25	N/A	N/A
Testperson 12	Tale inn, Skjerm ut	7,5	Mus inn, Skjerm ut	12
Testperson 13	Tale inn, Skjerm ut	8	Mus inn, Skjerm ut	17
Testperson 14	Tale inn, Skjerm ut	9,75	Mus inn, Skjerm ut	14,5
Testperson 15	Tale inn, Skjerm ut	8	N/A	N/A
Testperson 16	Tale inn, Skjerm ut	8,6	N/A	N/A
Testperson 17	Tale inn, Tale ut	51	N/A	N/A
Testperson 18	Tale inn, Skjerm ut	9,6	N/A	N/A

Kun testperson 17 benyttet tale som output – han hadde ikke mulighet for å se på skjerm. Han benyttet i gjennomsnitt 51 sekunder på enkle værøsk. Denne testpersonen var imidlertid blant de to som hadde til dels store problemer med å beherske tale som input, slik at ikke bare output løsningen er årsak til den langt høyere tidsbruken.

Den første oppgaven for testperson 11, oppgave 2 for testperson 12 og oppgave 3 for testperson 17 er utelatt fra utregning av gjennomsnittlig tidsbruk, da tiden ble forlenget grunnet feil talegjenkjenning. Oppgave 6 for testperson 14 er utelatt fra utregning av gjennomsnittlig tidsbruk, da tiden ble forlenget grunnet en systemsvikt som ga treghet i systemet.

1.6.2.7 Totalopplevelse

Testerne foretrakk å benytte tale input og skjerm output. Det ble opplevd som naturlig å kunne snakke inn til systemet, og som avlastende da antallet nødvendige klikk i grensesnittet reduseres;

“(…) det blir mindre trykking for å spørre enn å trykke rundt.” – Testperson 16

Taleinput ble opplevd som enklere enn identifisering av og interaksjon med det grafiske grensesnittet. Bruk av tale i slike grensesnitt vurderes derfor som positivt og var også den tidsmessige mest effektive løsningen.

Den nest mest effektive metoden var museløsning inn og skjermløsning ut. Denne metoden ble kun benyttet når testerne ble påtvunget å bruke museløsning som input – dvs. i oppgave 3 og 4.

1.6.3 Dyslektikere og personer med lese- og skrivevansker

Tabell 7: Oversikt over testpersoner med dysleksi/lese- og skrivevansker

	Alder	Beskrivelse av funksjonsproblematikk	Hjelpemidler i forhold til PC-bruk
Testperson 21	18	Dysleksi	Smartpen
Testperson 22	19	Dysleksi	Lingdys, Lingwrite, Daisy, cd-ord
Testperson 23	47	Dysleksi, skriveutfordringer - spesielt grammatikalsk.	Ingen hjelpemidler
Testperson 24	12	Dysleksi, problemer med rettskriving og å skrive pent.	LingDys (stavekontroll, ordbøker, ordfullføring, skjermleser og kunstig tale gjør at programmet gir omfattende og effektiv hjelp til både lesing og skriving) og LingRight(for å hjelp norskspråklige å skrive engelsk)
Testperson 25	38	Lese og skrivevansker	Ingen hjelpemidler
Testperson 26	38	Dysleksi	Ingen hjelpemidler
Testperson 27	20	Lese og skrivevansker	Ingen hjelpemidler
Testperson 28	15	Lese og skrivevansker	Ingen hjelpemidler
Testperson 29	19	Dysleksi	LingDys
Testperson 30	15	Dysleksi	LingDys

I denne gruppen har syv av ti testpersoner diagnosen dysleksi, hvorav fire bruker hjelpemidler for Dysleksi (for eksempel LingDys). De øvrige tre i gruppen har udiagnostiserte lese- og skrivevansker – som derfor også kan være relatert til andre vansker; for eksempel konsentrasjonsvansker eller generelle lærevansker. Aldersspennet i gruppen er fra 12 til 47 år.

1.6.3.1 Valg av input

Alle testpersonene i gruppen hadde muligheter for å benytte både tastatur og mus. De var imidlertid mest fortrolig med å bruke en standard museløsning, og tastaturet ble ikke sett på som en relevant input metode for denne gruppen.

På de frie oppgavene delte gruppen seg grovt sett i to leire; halvparten valgte tale som input over mus, mens den andre halvparten prefererte datamus over tale. Denne delingen var uavhengig av alder eller diagnostetype – og inntrykket var at de enten ubevisst valgte mus av vane eller bevisst valgte tale for å teste dette. Det kan likevel se ut som om flere velger tale ettersom de blir fortrolige med dette for enkle værøsk (oppgave 7), mens de foretrekker mus når oppgaven blir mer ukjent og kompleks (oppgave 8). Tabell 8 viser fordelingen av bruk av input metoder.

Tabell 8: Antall testere per benyttede inputmetoder

	Oppgave 1	Oppgave 2	Oppgave 3	Oppgave 4	Oppgave 5	Oppgave 6	Oppgave 7	Oppgave 8
Tale	5	5	-	-	10	10	7	3
Mus	5	5	10	10	-	-	3	7

1.6.3.2 Opplevelse av tale som input

“... Man må bli kjent med korrekt tale så den skal skjønn deg.” – Testperson X

Fire av de ti testerne opplever at talestyring er den mest effektive og behagelige input metoden; “... bra at man kunne snakke vanlig og ikke anstrenge seg for å snakke tydelig.” Fem testere sier at de foretrekker å bruke mus; “Med mus kan det kanskje bli litt lettere. Da kan du velge hva du vil, i stedet for at PC-en kanskje ikke forstår deg”. Den siste testeren er uklar; han er positiv til konseptet men måtte ha ekstra fokus ved tale, fordi han stammer. Han virket derfor usikker på hvilken interaksjonsmetode han foretrakk. Han valgte imidlertid talestyring i alle de frie oppgavene.

Selv om fem personer rapporterte å foretrekke mus, og kun fire tydelig sa at de foretrakk tale, viser observasjonene at seks personer i hovedsak benytter tale – dvs. at to testere benyttet noe annet i testene enn de sa at de foretrakk i intervjuene.

1.6.3.3 Valg av output

Alle testerne benyttet skjerm som output metode.

1.6.3.4 Opplevelse av det visuelle grensesnittet

Fire av testerne nevner at symboler, korte ord og tall gjør det enkelt å lese resultatet fra skjermen. Ni av testerne mener at de ville ha foretrukket å bruke tale dersom det var mye tekst å lese, eller teksten var mer komplisert. Testperson 23 sier at talen ville ha vært nyttig “Hvis man kan følge tekst på skjerm når det leses opp informasjon”, altså dersom:

- Highlighting følger tekst på skjerm når det leses opp informasjon.
- Man kan justere lesehastigheten til egen lesehastighet

To testere opplever at det visuelle grensesnittet ikke er brukervennlig for å gi input. De sier “... for mange knapper for (å velge) dag og tid.” og “... knappene er veldig like. Rammer i farger rundt knappegrupper ville ha gjort det lettere.”.

1.6.3.5 Tid brukt

I tabell 9 finner du gjennomsnittlig tidsbruk på enkle værøsk. Fem av de ti testerne slet med gjenkjenningsfeil som følge av en systemsvikt relatert til rekkefølgen på nøkkelord. Dette forlenget tidsbruk i følgende oppgaver, som ble ekskludert fra tidsberegninger: oppgave 2 for testperson 29, oppgave 4 for testperson 23, oppgave 5 for testpersoner 26, 27 og 29, oppgave 6 for testperson 29 og oppgave 7 for testperson 25.

Det totale gjennomsnittet ved tale input (skjerm output) for brukergruppen er på 10,4 sekunder.

Det totale gjennomsnittet ved mus input og skjerm output for brukergruppen er på 14,9 sekunder.

Tabell 9: Tidsbruk per testperson, relatert til interaksjonsmetoder

	Foretrukket metode	Gjennomsnittlig tid (sekunder)	Øvrige metoder	Gjennomsnittlig tid (sekunder)
Testperson 21	Tale inn, Skjerm ut	7,3	Mus inn, Skjerm ut	11,25

Testperson 22	Tale inn, Skjerm ut	7,5	Mus inn, Skjerm ut	15
Testperson 23	Tale inn, Skjerm ut	15,4	Mus inn, Skjerm ut	24
Testperson 24	Tale inn, Skjerm ut	7,8	Mus inn, Skjerm ut	14,5
Testperson 25	Tale inn, Skjerm ut	21,75	Mus inn, Skjerm ut	19
Testperson 26	Mus inn, Skjerm ut	14	Tale inn, Skjerm ut	7,5
Testperson 27	Mus inn, Skjerm ut	12,2	Tale inn, Skjerm ut	9,5
Testperson 28	Mus inn, Skjerm ut	13,6	Tale inn, Skjerm ut	10
Testperson 29	Mus inn, Skjerm ut	8,3	Tale inn, Skjerm ut	10
Testperson 30	Tale inn, Skjerm ut	7,3	Mus inn, Skjerm ut	17

Det er lite forskjell i tidsbruken benyttet for de som sier de prefererer musebruk versus de som sier de foretrekker tale som input. Uavhengig av testpersonenes subjektive oppfatninger og preferanser, virker bruk av tale input å være noen sekunder mer effektivt enn musebruk.

1.6.3.6 Totalopplevelse

Generelt var testerne positive til piloten, og til å ha taledialog som en del av grensesnittet. Flere fokuserte imidlertid på at det var fint å ha et *forenklet visuelt* grensesnitt – slik piloten hadde.

De aller fleste mente tale input løsningen var enkel å bruke. Et par av testerne virket ambivalente i forhold til løsningen. To testpersoner hadde etter vår vurdering vansker med å benytte løsningen.

På tross av at testene viser at bruk av tale er mer effektivt enn musebruk, virket det som om flere følte seg mer komfortable med bruk av mus enn tale og usikre på den nye interaksjonsmuligheten.

1.7. Oppsummering av hovedfunn

Det har vært en stor bredde i testdeltakerne, også internt i de ulike brukergruppene. Generelt kan vi si at datakunnskap, grad av funksjonshemming og erfaring med hjelpemiddelteknologi har vært svært varierende.

Synshemmede testere hadde svært differensierte måter å interagere med PC på. På tross av dette ble talemødaliteter i stor grad nyttiggjort som støtte i PC-interaksjon. Videre var talemødaliteter en kjempestor fordel for synshemmede med liten kunnskap om hjelpemiddelteknologier. Fra ikke å være i stand til å benytte grensesnittet, kunne testpersonene 2, 4 og 5 utføre værøøk like effektivt som synshemmede med høy datakompetanse. Det antas at synshemmede uten kjennskap til PC og hjelpemiddelteknologi vil ha svært stor nytte av taledialoger i grensesnitt.

Bevegelsehemmede testere var også differensierte, med ulike typer finmotoriske vansker og ulike grader av interaksjonsproblematikk. En god del av testerne var ikke i stand til å benytte standard museløsninger. Ved bruk av tale som input modalitet, ble imidlertid interaksjonen til testerne like effektiv uansett grad av håndkontroll, og like effektiv som seende testpersoner uten bevegelsehemming (for bevegelsehemmede var tiden benyttet ved tale som input og skjerm som output 8,7 sekunder i snitt, mens den testgruppen med lese- og skrivevansker var 10,4 i snitt). Alle

bevegelseshemmede testere unntatt en ga utrykk for at tale input forenkler interaksjonen, er intuitivt, effektivt og mindre anstrengende enn å benytte enn alternative inputmetoder.

Tale som input viste seg i testoppgavene å være tidsmessig mer effektivt enn både tastatur og mus for de aller fleste testere. Tale ble her nesten alltid målt til noen sekunder mer effektivt enn mus, nokså uavhengig av testpersoners funksjonsnedsettelse, preferanser og subjektive oppfatninger. Sammenliknet med tastaturbruk ga tale tidsbesparelser på i overkant av 10 sekunder, og ble også vurdert som energibesparende.

For musebrukere ble det imidlertid tydelig at ikke bare effektivitet spiller en rolle for brukervennlighet og preferanse. Testere vant med standard museløsninger splittet seg i to jevnstore sub-grupper, der den ene ønsket bruk av tale mens den andre foretrakk museløsninger. Det var ingen klare skillelinjer mellom de to sub-gruppene, men det kunne virke som om flere testere følte seg komfortable med mus og kjente modaliteter, og følte seg usikre på kontrollen ved talekommandoer.

Tilbakemeldinger indikerer at tale output er noe dårlig tilpasset for seende brukergrupper, og bør gjøres mer naturlig og gis justeringsmuligheter. For seende er det mest effektivt å bruke skjerm som output metode.

1.8.Konklusjon

Testene viser at taleinput er en enkel og effektiv metode og medfører en belastningslettelse. Det ble identifisert to grupper med store vansker i forhold til PC-interaksjon, som hadde kjempenytte av talemodaliteter – og spesielt tale som input modalitet. Den ene er bevegelseshemmede som ikke kan benytte standard museløsninger, den andre er synshemmede med lav tastaturkompetanse. For disse gruppene gjør talemodalitetene grensesnittet tilgjengelig, og bidrar således til nødvendig redundans.

Resultatene indikerer at bruk av tale input kan være noe raskere enn bruk av datamus. Dette er svært interessant, og det hadde vært spennende å gjennomføre testene med en kontrollgruppe ikke-funksjonshemmede brukere for å undersøke effektivitetsspørsmålet nøyere, med mer komplekse oppgaver, i større skala og i en mer homogen gruppe, og foreta mer nøyaktige observasjoner av tidsbruk.

Vedlegg A: Nettbasert multimodalt dialogsystem for værvarslingsinformasjon

Dette er første skisse av en testpilot i SMUDI prosjektet på et nettbasert multimodalt dialogsystem. Hele systemet skal kunne benyttes kun ved tale/audio info og kun ved trykk/visuell info.

Som nevnt i SMUDI prosjektsøknadens del 2 "Kunnskaps- og teknologifronten", finnes det relevante standardiseringsarbeider blant annet i W3C. Arbeidsgruppen "Multimodal Interaction Activity" (MMI) (<http://www.w3.org/2002/mmi/>) utvikler spesifikasjoner som skal gjøre det mulig å bruke multimodal interaksjon med nettleseren. En viktig del er EMMA (Extensible MultiModal Annotation markup language) – et "markup language" som skal transkribere og forstå brukernes ytring; for eksempel kombinere ord de sier med trykk på en skjerm.

Jf. W3C om lyttemodus, anbefales dialogsystemet å baseres enten på "evig lytting" eller på såkalt "push to activate". Ved "evig lytting", lytter systemet konstant, typisk etter et bestemt nøkkelord som så trigger input av talestrøm. Denne måten å lytte på er tenkt benyttet i dette dialogsystemet. Ved "push to activate" trykker brukeren på en bryter/knapp i forkant av tale/kommando, og deretter gir taleinput til systemet. Systemet detekterer selv når brukeren er ferdig med å gi taleinput. Systemet kan selv time ut dersom ingen input kommer inn.

Systemet kan vekkes til å starte opp ved input via tastatur/mus/bryter/nøkkelord. I en stasjonær setting, kan også sensorer brukes for å oppdage tilstedeværelse av bruker.

Systemstatus bør alltid presenteres til brukeren. Brukeren må få en feedback på at systemet venter på input, prosesserer data eller har fått en systemfeil. Under er W3C forslag til statusmeddeling. Når det gjelder ikoner, må form brukes i tillegg til farge grunnet fargeblindhet. Forslag her er en "stående grønn person" (klar) og en "gående rød person" (jobber). Dialogen vil alltid auditivt indikere at systemet venter input, men busy og error er relevante statuser å presentere til brukeren.

Table 4: Suggested indicators for the current system status			
Mode	Idle	Busy	Error
Text	"Ready for next input"	"Processing, please wait"	Explanation for the cause of the error and how to fix it
Icons	Green*	Red*	Blinking "danger" icon
Audio	Silence	Sounds of a clicking clock or a percolating coffee pot	Emergency vehicle siren

Dialogsystemets funksjonalitet

Dialogsystemet vil dra nytte av informasjon fra YR for å tilby følgende tjenester:

Værvarsel - Type	Beskrivelse	Fremstillinger
Kort oversikt	Dagens og neste dags varsel: formiddag – dagtid – kveld (dagens f.o.m. nåværende varsel)	Tabell med opplesing Dialog
Time-for-time	Varsel for de neste 24 timer	Tabell med opplesing Dialog Meteogram (24/48 timer)
Helgevarsel	Varsel for fredag til søndag: natt – formiddag – dagtid – kveld (fredagens f.o.m. nåværende varsel)	Tabell med opplesing Dialog
Langtidsvarsel	Dagens og kommende 7 dagers varsel: middagsvarsel per dag	Tabell med opplesing Dialog
Direkte dialog	Dialog utenom varslene	Dialog – spørsmål/svar

Direkte dialog

Tjeneste	Auditiv dialog	Visuell fremstilling
Velkommen	System: Velkommen til norsk værsøk!	Tekst " Velkommen til norsk værsøk!"
Introduserende dialog	System: Si meg hvilken by du ønsker varsel fra, og hva slags type varsel du ønsker! Ved feil, si FEIL. (pause) System lytter etter tilgjengelige byer, tilgjengelige varseltyper, ukedager + i dag + i morgen og tidspunkt + formiddag + ettermiddag + kveld + dagtid, og FEIL. Ved timeout/ukjent input: Værsøk-sted. IKKE-STED: Dersom sted ikke angis, til værsøk-sted. FEIL: Velkommen... Ellers: Gjenkjent stedsinput settes tekstlig i søkefelt, og systemet bekrefter:	1. Øverst: Søkefelt med søkeknapp. Hjelpetekst foran søkefelt: "Si eller skriv inn hvor du ønsker varsel fra." Mouse-over hjelpetekst til søkefelt: "Søk etter stedsnavn". Label søkeknapp: "SØK" Godt synlig musepeker i farge/størrelse, som endres til innsettingsmarkør over søkefelt. 2. Sentrert overskrift: "Si meg hvor du ønsker varsel fra, og hva slags type varsel du

	<p>STED: Går til Værvarsel - type</p> <p>STED+TYPE: Registrert varselønske leses.</p> <p>STED+DATO: Du ønsker varsel for <sted>, på <dag>. (Langtidsvarselet for dagen leses opp).</p> <p>STED+TID (kun tid, ikke dato -> Dersom kun sted og time spesifiseres, antar man dagen i dag som dato): Du har valgt varsel for <sted>, i dag kl <time>: ... (Time for time varsel for angitt time leses.)</p> <p>STED+TID2 (kun tidspunkt, ikke dato -> antar dagen i dag): Du har valgt varsel for <sted>, i formiddag/dag/kveld: ... (Detaljert langtidsvarsel for angitt tidspunkt leses.)</p> <p>STED+TYPE+DATO: Du har valgt <varseltype> for <sted>, <dato/dag>: ... (Angitt varseltype for dagen leses opp).</p> <p>STED+DATO+TID: Du har valgt varsel for <sted>, <dag> kl <time>: ... (Time for time varsel for angitt time leses.)</p> <p>STED+DATO+TID+TYPE: Du har valgt <varseltype> for <sted><dag><tid>: ... (Detaljert langtidsvarsel/time-for-time leses.)</p>	<p>ønsker!"</p> <p>3. Under overskriften: "Ved eventuelle feil, si FEIL."</p> <p>4. Innhold: Evtnt. tabeller over mest brukte/sist brukte steder. Tekst: "Hvilken type varsel ønsker du: Kort oversikt, Time for time, helgevarsel eller langtidsvarsel? Ønsker du en bestemt dag og tid, si f.eks. 'onsdag kl 12'."</p> <p>5. Nederst: WYSIWYG skjermtastatur. Mulighet for å minimere tastatur ved å klikke på – tegn. Mulig å maksimere igjen ved + tegn. Mulig å gjøre mindre/større ved å dra i det?</p>
Værøk – sted	<p>Hvilken by i Norge ønsker du værvarsel fra?</p> <p>(Det er 96 byer i Norge. Systemet kan enten lytte etter all 96 eller kun et utvalg, som da må ramses opp, eks: Lillehammer, Oslo, Arendal, Stavanger, Bergen, Trondheim, Bodø, Tromsø eller Svalbard?)</p> <p>Systemet lytter etter tilgjengelige byer og tilgjengelige varseltyper, ukedager + i dag + i morgen, tidspunkt og FEIL.</p> <p>Time-out: Velkommen + direkte dialog.</p> <p>FEIL: direkte dialog</p> <p>Ukjent input: gjenta Værøk-sted fra topp.</p> <p>Ellers: gjenkjent input settes tekstlig i søkefelt, og til værvarsel - type.</p>	Som for direkte dialog.
Værvarsel – type,	<p>System: Du har valgt varsel for <sted>. Hvilken type varsel ønsker du: Kort oversikt, Time for time, helgevarsel eller langtidsvarsel? Ønsker du en bestemt dag og tid, si for eksempel "onsdag kl 12". Ved</p>	Som for direkte dialog, men med oppgitt sted i søkefeltet.

	<p>feil, si FEIL.</p> <p>Systemet lytter etter Kort oversikt, Time for time, helg/helgevarsel og langtids/langtidsvarsel, ukedager + i dag + i morgen, tidspunkt, og FEIL.</p> <p>Time-out: Velkommen + direkte dialog.</p> <p>FEIL: Til direkte dialog.</p> <p>Ukjent input: gjenta Værvarsel-type fra topp.</p> <p>Gjenkjennbare type/dag/time kombinasjoner: som for direkte dialog.</p> <p>Ellers: registrert varselønske leses.</p>	
--	--	--

Kort oversiktsvarsel

Tjeneste	Auditivt	Visuelt
Tabell med opplesing	<p>Kort oversiktsvarsel (for <sted>).</p> <p>Informasjonen i tabellen leses opp som følger: ”<dato> <tidspunkt> <tekstlig beskrivelse> <grader> <...>.</p> <p>Du har nå hørt oversiktsvarsel for <sted>.</p> <p>Ett earcone: ikonet for vær.</p> <p>Nedbør (regn/snø): prikkerytme</p> <p>Solskinn, uavhengig av mengde: dempet fuglekvitte</p> <p>Overskyet: svak suselyd</p>	<p>Overskrift: Kort oversiktsvarsel (for <sted>).</p> <p>I Innholdsdel: Tabell som på yr.no i dag.</p> <p>I alle varseltyper, vises faneblader med de ulike typer varsel, jmf. yr.no per i dag. Overskrift og tabell vises rett under søkefelt som forblir liggende øverst.</p>
Dialog	<p>Systemet lytter etter FEIL for å avbryte varsel.</p> <p>FEIL/ferdig opplesing: til Velkommen.</p>	

Time for time

Tjeneste	Auditivt	Visuelt
Tabell med opplesing	<p>Time for time varsel (for <sted>)</p> <p>Informasjonen i tabellen leses opp som følger: ”<overskrift dag 1> <time> <beskrivelse ikon><nedbør> <grader><vind og vindretning><time><...><overskrift dag 2><time><...>.</p> <p>Du har nå hørt time for time varsel for <sted>.</p> <p>Ett earcone: ikonet for vær.</p>	<p>Overskrift: Værvarsel time for time (for <sted>).</p> <p>I Innholdsdel: Tabell som på yr.no i dag.</p> <p>I alle varseltyper, vises faneblader med de ulike typer varsel, jmf. yr.no per i dag. Overskrift og tabell vises rett</p>

	Nedbør (regn/snø): prikkerytme? Solskinn, uavhengig av mengde: dempet fuglekvitte? Overskyet: svak suselyd?	under søkefelt som forblir liggende øverst.
Dialog	Systemet lytter etter FEIL for å avbryte varsel. FEIL/ferdig opplesing: Velkommen...	
Meteogram	-	Meteogram som på yr.no Vises mellom overskrift og tabell, dvs. øverst i innholdsdel.

Helgevarsel

Tjeneste	Auditivt	Visuelt
Tabell med opplesing	Helgevarsel (for <sted>) Informasjonen i tabellen leses opp som følger: " <dato> <tidspunkt> <ikon beskrivelse><nedbør> <grader><vind><vindretning> <...>. Du har nå hørt helgevarsel for <sted>. Ett earcone: ikonet for vær. Nedbør (regn/snø): prikkerytme Solskinn, uavhengig av mengde: dempet fuglekvitte Overskyet: svak suselyd	Overskrift: Helgevarsel (for <sted>). Innholdsdel: Tabell som på yr.no i dag. I alle varseltyper, vises faneblader med de ulike typer varsel, jmf. yr.no per i dag. Overskrift og tabell vises rett under søkefelt som forblir liggende øverst.
Dialog	Systemet lytter etter FEIL for å avbryte varsel. FEIL/ferdig opplesing: Velkommen...	

Langtidsvarsel

Tjeneste	Auditivt	Visuelt
Tabell med opplesing	Langtidsvarsel (for <sted>) de neste 7 dager. Informasjonen i tabellen leses opp som følger: " <dato> <tekstlig beskrivelse><grader> <...>. Du har nå hørt langtidsvarsel for <sted>. Ett earcone: ikonet for vær. Nedbør (regn/snø): prikkerytme Solskinn, uavhengig av mengde: dempet fuglekvitte Overskyet: svak suselyd	Innholdsdel: Tabell som på yr.no i dag. I alle varseltyper, vises faneblader med de ulike typer varsel, jmf. yr.no per i dag. Overskrift og tabell vises rett under søkefelt som forblir liggende øverst.
Dialog	Systemet lytter etter FEIL for å avbryte varsel. FEIL/ferdig opplesing: Velkommen...	

Vedlegg B: Teknisk spesifikasjon og dialogeksempel

Dette vedlegget er levert av Veridict, Sverige og kan derfor inneholde noen språklige misligheter.

The application consists of a stand-alone software application with a graphical user interface, and a voice user interface. The voice user interface will consist of spoken input by speech recognition and spoken output by speech synthesis. The application demonstrates multimodal interaction with the online service YR.NO.

Users will be able to ask simple questions (queries) to the system, either by using voice or by mouse and get multimodal responses based on current actual data on the YR.no site. The terminology will be constrained to a very limited set of queries and a limited set of ten different cities. The graphical user interface (GUI) and the voice user interface (VUI) will be in synchrony, i.e., any input either by mouse in GUI or by voice in the VUI will be reflected in the opposite interface.

Below is an example of a possible dialog:

System: Welcome to YR.no! If you say 'What can I say' you will be given some examples. Tell me what city you want to select?

User: Choose Oslo.

System: Downloading the current Oslo weather. (Pause) Done!

User : When does the sun rise?

System: The sun will rise at 9:29 and set at 15:19

User: What can I say?

System: For example you could say 'Week summary' for an overview of the week, or you can say 'What's the tomorrow's weather' or more specific 'What's the weather Friday afternoon'.

User: Tomorrow's weather

System: Thursday early morning. Partly cloudy. -12 degrees. Strong wind, Northern 3 meter

System: Thursday morning. Partly cloudy. -10 degrees. Strong wind, Northern 3 meter

System: Thursday afternoon. Snow. -13 degrees. Strong wind, Northern 1 meter

System: Thursday evening. Partly cloudy. -15 degrees. Light winds, Northern 2 meter

User: What's the weather Friday afternoon'

System: Friday afternoon'. Sun. -5 degrees. Breeze, South 1 meter

User: Week summary

System: Østlandet torsdag og fredag: Skiftende bris. Pent vær, men lokal tåke eller lave tåkeskyer.

System: Østlandet laurdag: Skiftende bris, men for det meste nordlig. Stort sett pent vær. Lokal tåke i lavlandet.

System: Østafjells søndag: Bris omkring øst, på kysten sørøstlig opp i kuling, fra om ettermiddageng dreierende sørvestlig og minkende på kysten av Telemark og Agder. Snø av og til, vesentlig sør for Lillehammer, mest nedbør fra Vestfold til østlige deler av Agder. Etter hvert stigende temperatur.

System: Sør-Norge måndag og tysdag: Skiftende bris, på kysten av Vestlandet periodevis nordlig av og til kuling. Periodevis enkelte snøbyger på Vestlandet og etter hvert i Trøndelag. Ellers for det meste oppholdsvær. Litt lavere temperatur de fleste steder.

System: Norge og Spitsbergen onsdag til laurdag: På Spitsbergen skiftende vind, kortvarig kuling utsatte steder. Av og til litt snø. Omtrent uendret temperatur. I Nord Norge ventes først i perioden sannsynligvis mest rolige vindforhold og mulighet for enkelte snøbyger på kysten. I slutten av perioden kan et lavtrykk vestfra berøre Nordland med vind og nedbør. Også Sør-Norge kan berøres av et par lavtrykk som i løpet av perioden kommer

vestfra. Det er knyttet betydelig usikkerhet til lavtryksbanene. Sannsynligheten for perioder med kuling og nedbør er størst på Vestlandet, i kortere perioder kan vind og nedbør også passere i Trøndelag og østafjells.

Vedlegg C: Informasjons- og samtykkeerklæring

Forespørsel om deltakelse i forskningsprosjektet:
SMUDI – Stemmestyring i Multimodal Dialog

Dette er et spørsmål til deg om å delta i en forskningsstudie ledet av MediaLT. Deltakerne skal gjennomføre noen oppgaver og blir intervjuet om ulike måter de kan anvende teknologi på, for å bruke tjenester på nettet.

Bakgrunn

Bruken av tjenester på nettet øker kraftig. Dessverre er de ikke like lette å bruke for alle. Dårlig utforming kan gjøre det vanskelig for personer med funksjonsnedsettelse å bruke disse tjenestene.

I SMUDI-prosjektet forsker vi blant annet på hvordan tale kan brukes i kombinasjon med tastatur, mus og skjerm for at tjenester på nettet i så stor utstrekning som mulig kan brukes av så mange som mulig. Som en del av prosjektet, har vi utviklet en demonstrator for å teste ut dette. For å få best mulige svar på dette forskningsspørsmålet, må vi prøve ut demonstratoren sammen med reelle personer. Vi ønsker derfor å invitere deg til å være med å prøve ut den nye løsningen.

Praktisk informasjon for deltakere

- Sammen med forskere fra MediaLT skal du gjennomføre noen praktiske oppgaver av demonstratoren.
- Du vil også bli stilt korte spørsmål om alder, hva slags funksjonsutfordringer du har, hva slags tekniske hjelpemidler du eventuelt bruker og hvordan du opplevde bruken av demonstratoren.
- Utprøvingen vil ta ca. 1 time.
- Du vil være med i trekningen av fem gavekort på tusen kroner som takk for at du deltar.
- Vi avtaler tid og sted for utprøvingen som passer for deg. Dette kan være hjemme hos deg, på skole/arbeidsplass, i våre lokaler i Oslo eller et annet egnet sted.
- Vi stiller med datautstyr for å utføre utprøvingen. Ved spesielle behov, og dersom dette ansees hensiktsmessig, kan egne hjelpemidler anvendes for å gjennomføre utprøvingen.

Prosjektet

SMUDI-prosjektet er et samarbeid mellom en rekke norske institusjoner, bedrifter og brukerorganisasjoner. Prosjektet mottar finansiering fra Norges forskningsråd og Kunnskapsdepartementet. Du kan lese mer om SMUDI på www.smudi.no.

Personvern

Deltakelse er frivillig. All innsamlet informasjon vil bli anonymisert og behandlet konfidensielt. Alle notater slettes ved prosjektslutt, eller senest juni 2012. Ingen utenfor prosjektet vil ha tilgang til den informasjonen du gir. Informasjonen vil bare bli brukt som grunnlag for forbedring av tjenester på nettet, samt til rapporter og vitenskapelige artikler og foredrag. Det vil ikke være mulig å identifisere enkeltpersoner.

Du kan trekke deg fra prosjektet og/eller kreve personopplysningene som er gitt, slettes når som helst uten å oppgi grunn.

Før intervju og utprøving begynner, bes deltageren om å samtykke ved å undertegne på at man har lest og forstått informasjonen, og at man ønsker å delta.

Kontakt undertegnede dersom du har spørsmål rundt deltakelse i prosjektet. Takk for at du ønsker å delta!

Vennlig hilsen

Morten Tollefsen, forskningsleder, MediaLT

E-post: morten@medialt.no

Telefon: 90 89 93 05

Samtykke

Jeg har lest og forstått informasjonen over, og gir mitt samtykke til å delta i undersøkelsen.

Sted og dato

Signatur

Vedlegg D: Observasjonsskjema og intervjuguide

Navn: _____ Tlf: _____
Alder: _____ e-post: _____
Bosted: _____
Hjelpemidler: _____
Funksjonshemning: _____

Oppgave nr:	Input	Output	Tid brukt
1			
2			
3			
4			
5			
6			
7			
8			

Forklaring til koder:

Input – TA=tale, TS=tastatur, MU=mus

Output – T=tale, S=skjerm, L=leselist

Fritekstnotater (angi evt. spesifikt oppgavenummer)

Intervju

Intervjuene søker å måle brukernes egne **opplevelser** av det multimodale grensesnittet, og **hvilke interaksjonsmuligheter** som foretrekkes.

1: Hvilken input metode og output metode opplevde du som mest effektiv og mest behagelig å bruke?

2: Synes du talegjenkjenningen fungerte tilfredsstillende?

3: Var det noe du syntes var spesielt bra eller spesielt dårlig med systemet?

Positivt:

Negativt:

4: Vil slike systemer gjøre det enklere for deg å få tilgang til informasjon fra internett? Begrunn svaret.

5: Kan du si noe om når du foretrekker å bruke ulike måter å gi input og få output på?
(Vi forklarer her at noen kan foretrekke å f. eks. lese telefonnummer med leselist, men å få lange tekster opplest med tale; at noen vil se grafiske symboler for været, men få tekst lest opp; at noen vil gi kommandoer med tale, men selv se resultater på skjerm osv..)