

Indholdsfortegnelse

Dineline konceptet.....	2
HCI- tale som in- og output.....	4
Talesyntese	4
Text-to-speech.....	5
VoiceXML.....	5
Interaktivitet og lyd.....	8
Sprog.....	10
Konstruktionen	12
Dineline – musikbrug	15
Anvendelse og omfang.....	15
Musikkens funktion	15
Lydlogo.....	16
Vignetter.....	17
Ophavsret	22
Afrunding.....	22

Dette projekt drejer sig om Voice XML og giver et eksempel på, hvordan man kan udnytte denne teknologi i praksis.

Vi har udviklet et koncept, som vi kalder Dinline, som er en interaktiv teletjeneste, der i realtime guider brugeren igennem tilberedningen af et måltid.

Vi har først og fremmest fremstillet et scenarie, som eksemplificerer, hvordan tjenesten fungerer, og hvordan den kan tage sig ud i henhold til en bestemt målgruppe. Desuden har vi forsøgt at redegøre for den teknologi, der ville skulle impliceres i et sådant produkt, og hvad konceptet overordnet indebærer af hensyn til interaktion, sprog m.m.

Vi har også prøvet at perspektivere projektet og sat konceptet i sammenhæng med teknologiske muligheder samt nogle økonomiske faktorer, og overvejet hvilke muligheder for anvendelse idéen indeholder.

Vores scenarie er kun ét ud af mange mulige, det overordnede koncept taget i betragtning, men vi har i dette tekstdokument søgt at fokusere på vores egen produktion for derved bedre at kunne sandsynliggøre og argumentere for konceptet.

I det følgende vil vi derfor beskrive overordnede forhold, der har betydning for Dinline-konceptet, og vi vil redegøre for vores specifikke produkt, samt hvilke overvejelser der ligger bag resultatet.

Dinline konceptet

Overordnet set er Dinline-konceptet en teletjeneste, der i princippet kan tilpasses mange 'gør det selv' aktiviteter, som eksempelvis vindueskitning eller vedligeholdelse af bilen. I vores eksempel er der tale om en interaktiv kokebog, der trin for trin guider brugeren gennem en proces, i dette tilfælde madlavning.

Udviklingen af vores koncept er i høj grad foregået med udgangspunkt i en teknologi, der gør det muligt netop at kunne betjene et computersystem ved hjælp af stemmen, da man for det første skal bruge sine hænder og øjne til madlavningen, og en computer i traditionel forstand, med skærm, tastatur og mus, næppe er den mest hensigtsmæssige sparringspartner i et køkkenmiljø, der er præget af fedtede fingre og stegeos.

Vores interaktive kokebog har generelt en meget bred målgruppe (stort set alle der gerne vil lave mad, og som ikke er alt for teknologi-forskrækkede). En bred målgruppe giver plads til meget vidde rammer for et koncept, hvilket selvfølgelig skaber en vis frihed, men som også kan gøre det svært ”at ramme

plet”. I introduktionen til kokebogen, inden valg af menu, forsøger vi at indsnævre vores målgruppe ved at fastsætte prisniveau, samt bede brugeren om at definere sig selv som enten ”begynder, let øvet, øvet” eller ”avanceret”. I vores demo-udgave er brugeren således på begynder niveau og vil gerne lave noget, der er billigt.

Vi forestiller os, at det er en dreng, der lige er flyttet hjemmefra, som lever af SU og derfor kun handler i Netto. Resten af systemet er derfor indstillet på en sådan bruger, både hvad angår valg af musik, speak, underholdning og menuvalg. Havde brugeren sagt at han/hun var ”avanceret” og ville lave noget dyrt, havde forløbet været markant anderledes. (Man ville fx ikke sige, at blusset skulle skrues ned, ”fx på 2”).

Med DineLine bliver det nu muligt for selv den mindst trænede køkkenakrobat, at lave en god middag til enhver lejlighed. DineLine er et koncept der kan guide en bruger gennem madlavningen, uanset hvilket gastronomisk niveau som han eller hun måtte befinde sig på.

Ved at ringe til DineLine, får brugeren adgang til en række målgruppetilpassede muligheder for at lave en middag, som passer til situationen. Der er i dette eksempel en lettere ironisk og humoristisk tone, både i guidens stemme, men også i valget af vignetter, som fx temaet fra ”Dødens gab” til valgmuligheden ’svigermor kommer’, der giver hele forløbet karakter af at være lettere satirisk. Denne tone kan ændres, alt efter hvilken målgruppe producenten stiler efter.

Konceptet for DineLine er baseret på, at fx dagligvarebutikker anvender konceptet i deres eget regi. Således kunne det hedde Nettos DineLine, hvorigennem man så kunne få opskrifter, baseret på fx ugens tilbud. Man kunne også forestille sig at et ungdomsblad, som fx ”Tjeck”, kunne have et samarbejde med en annoncør, da de begge har interesser i målgruppen. Det kunne være gratis for læserne at ringe ind på ”Tjecks DineLine”, og blive guidet igennem den perfekte score-middag, hvor indkøbssedlen leveres af annoncøren. Denne indkøbsseddel kunne fx sendes til brugerens mobiltelefon, og ville være en liste over de ingredienser, der skal anvendes i den pågældende ret, men den kan jo naturligvis også indeholde ugens tilbud (Selv om disse ikke nødvendigvis indgår i opskriften).

Vi forestiller os desuden, at man (brugerne eller fx ’ugens kok’) kan indsende opskrifter som e-mail, og disse blot sorteres for spam (så der ikke kommer opskrifter på ostegratineret havregød), og derefter blot lægges ind i databasen, hvor computeren ved hjælp af text-to-speech (TTS) gør opskriften tilgængelig for brugerne.

Med de nuværende telefontakster er konceptet nok mindre interessant i Danmark, med mindre at det fx baseres på gratis numre, hvor service- (Dineline) udbyderen, har købt x-mængde dataoverførsel af et teleselskab. En anden mulighed for finansiering kunne være at serviceudbyderen betaler for dataoverførelsen, og får det finansieret gennem reklamer, og/eller det forventede mersalg som følge af den medfølgende indkøbsseddel.

HCI- tale som in- og output

Kommunikation med computere ved hjælp af tale kan, som nævnt, være til stor hjælp i situationer, hvor det er nødvendigt at have hænder og øjne fri til andre opgaver end at kommunikere med computeren.

Ved hjælp af VoiceXML (Extensible Markup Language) vil det være muligt at kunne betjene en server/computer på afstand via fx en almindelig telefon sat på *håndfri mode*. Og talesyntesen gør at man ikke behøver at holde øje med en skærm.

Talesyntese

I dag er talesyntese hovedsageligt baseret på software, hvor det tidligere også kunne være ekstra hardware. Talesyntese benyttes til at få læst computertekst højt. Talesyntesen analyserer teksten og læser den højt efter de regler som gælder for det aktuelle sprog.

Den mest anvendte talesyntese er difonbaseret. I en difonbaseret talesyntese til dansk, består computerens *stemme* af 2600 præ-indspillede lydstumper, som kaldes for difoner. En difon består af slutningen af et fonem, og begyndelsen af det efterfølgende fonem.¹ Og det er disse difoner der så sammensættes til ord og sætninger af talesyntesen.²

Fordelen ved den difonbaserede talesyntese i forhold til en syntese der benytter sig af samlede ord er, at den difonbaserede talesyntese altid vil kunne udtale et skrevet ord, selvom det er et ualmindeligt (og derfor ikke indlæst) ord eller det er stavet forkert.

Talesyntese anvendes allerede i programmer, der læser tekst op for blinde og svagtseende, samt fx hos Teledanmarks nummeroplysning. Der er stadig lidt problemer med at få den helt rigtige/naturlige udtale af nogle ord, da ordene er sammensatte af enkelte stumper, og derfor ikke har den naturlige tales *flow*.

¹ <http://kom.auc.dk/~tb/articles/fontal97.pdf>

² <http://cpk.auc.dk/syntese/>

Vores Dinline computer skal ved hjælp af *Synthesis by rule*³, kunne generere tale ved hjælp af en difonbaseret talesyntese. Dette vil være det mest hensigtsmæssige valg, da det ellers ville være nødvendigt at indtale alle opskrifter der bliver indsendt som e-mail, og samle disse. Det er altså, langt det mest økonomiske at anvende en difonbaseret talesyntese, frem for at anvende samlede ord. Dog vil standard tale, som fx velkomst og introduktion være pre-indspillede samples, da disse vil være ens uanset hvilke valg man ellers foretager.

Text-to-speech

Som beskrevet i vores konceptbeskrivelse, er det vores tanke at det skal være muligt at indsende opskrifter, scorehistorier og andet til Dinline. For at denne funktion ikke skal blive en for stor byrde for Dinline-udbyderen, skal *indlæsningen* af de tilsendte opskrifter og historier kunne foregå automatisk. Dette kan lade sig gøre med et *tekst-to-speech* program.

VoiceXML

VoiceXML er en åben standard for beskrivelse af stemmestyrede dialoger baseret på W3C-standardens XML.

VoiceXML formatet minder om HTML, men er bare til *voice-browsers*, som fx kan benyttes via en telefon. Det er designet med henblik på at skabe lydaserede dialoger baseret på output som *text-to-speech*, talesyntese og digitaliseret lyd (herunder også præindspillede ord), og inputtet som automatisk talegenkendelse (ASR), interaktiv talerespons (IVR), *Nummertoner* (DTMF) og optagelse af talt input.⁴

Ved hjælp af VoiceXML, er det muligt at skabe en dialog mellem en service og en person ved hjælp af stemmestyling. VoiceXML er browseren der kobler sig til den eksisterende telefoni- og internetinfrastruktur. Princippet har været kendt længe, som fx valg ved brug af nummertoner fra trykknapp telefoner. "Tryk 1 for vejledning, tryk 2 for personlig betjening..."

³ Synthesis by rule: talesyntese baseret på samlede ord eller ord genereret af fonemer/difoner

⁴ <http://www.w3.org/TR/speech-synthesis/>

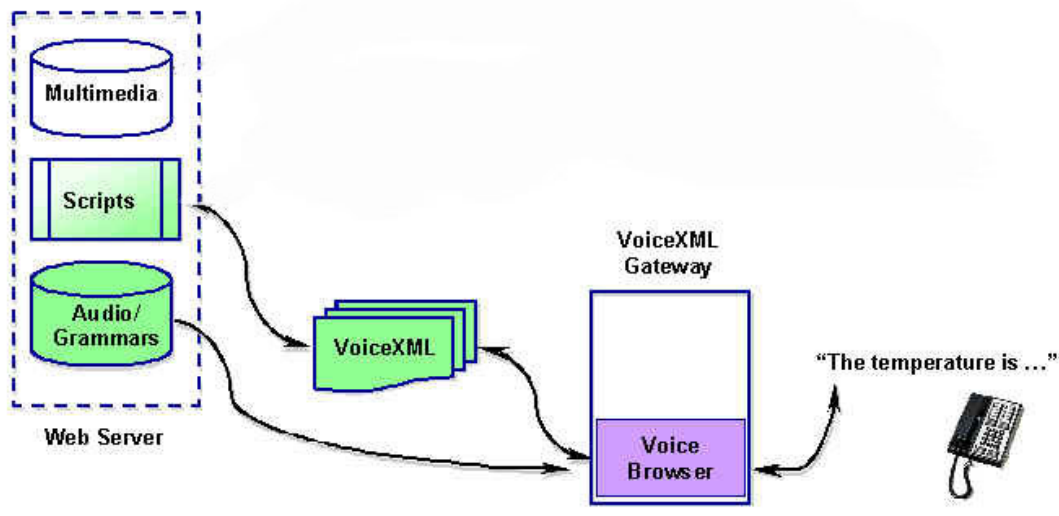


Fig. 1⁵

Fra telefonen er det muligt at ringe op til den ønskede server, i vores tilfælde *Dineline* serveren, og få læste opskrifter op. Brugerens stemme transformeres til VoiceXML, som får serveren til at sende de relevante oplysninger tilbage, genereret som syntetisk tale.

En dialog kunne forløbe således:

Computer:

Vælg prisniveau: Billigt; Smådyrt; Dyrt.

Bruger:

Gratis

Computer:

Jeg forstod ikke hvad du sagde.

(en platform-specifik default-besked.)

Computer:

Vælg prisniveau: Billigt; Smådyrt; Dyrt.

Bruger:

Billigt

Computer:

(går til <http://www.billigt.example/start.vxml>)

I VoiceXML ville det se således ud:

```
<?xml version="1.0"?>
```

```
<vxml version="2.0">
```

```
<menu>
```

```
  <prompt>
```

```
    Vælg prisniveau: <enumerate/>
```

⁵ <http://www.voicexml.org/tutorials/intro2.html>

```

</prompt>
<choice next="http://www.billigt.example/start.vxml">
  Billigt
</choice>
<choice next="http://www.smådyrt.example/intro.vxml">
  Smådyrt
</choice>
<choice next="http://www.dyrt.example/news.vxml">
  Dyrt
</choice>
<noinput>Vælg prisniveau <enumerate/></noinput>
</menu>
</vxml>6

```

VoiceXML er en åben standard, hvilket betyder at man ikke er afhængig af fx en bestemt platform eller support fra en specifik softwareproducent. Desuden er det som nævnt baseret på XML, hvilket gør det let tilgængeligt for programmører der i forvejen er bekendte med XML. VoiceXML er således det format der lader til at have det største potentiale, og det er allerede i anvendelse.⁷

Der arbejdes på at udvikle en slags *add-on* til VoiceXML kaldet SSML (Speech Synthesis Markup Language). Formålet med SSML er at give udviklere af indhold til en *voice-browser* en standard indenfor hvilken man kan kontrollere forskellige aspekter af tale, som fx udtale, lydniveau og pitch.⁸ SSML er dog langt fra færdigudviklet, men det vil til sin tid være en naturlig del af arbejdet med at skabe indhold baseret på VoiceXML.

Talesyntese	Den automatiske generering af tale-output fra input der fx kan være tekst.
Text-To-Speech	Processen hvor tekst-input bliver genereret automatisk til tale-output.
Voice Browser	Kan <i>forstå</i> (voice) markup language, som fx VoiceXML, og er i stand til at generere tale og/eller forstå tale input.

⁶ Eksemplet er hentet fra: <http://www.w3.org/TR/2001/WD-voicexml20-20011023/#dml1.3.1> og tilpasset vores koncept

⁷ Artikel fra teknologisk institut om VoiceXML:

<http://www.cit.teknologisk.dk/omcit/biblioteket/artikler/xml/voicexml.asp>

⁸ <http://www.w3.org/TR/speech-synthesis/>

Interaktivitet og lyd

I det følgende vil vi beskrive hvordan HCI (Human Computer Interaktion) med lyd som input og output, kan anvendes. Vi tager naturligvis udgangspunkt i vores Dinline.

At kunne betjene teknologi ved hjælp af lyde, er ikke et nyt fænomen, men det er heller ikke udviklet til sit fulde potentiale. Der findes mange miljøer, hvor den traditionelle computer, med knapper og tastatur, ikke ville overleve længe. Det kunne være på det olieindsmurte autoværksted eller i det fedtede køkken. Køkkenet er et værksted for madlavning, hvor man har hænderne i varerne, og ofte vil have gavn af at have sin opmærksomhed rettet mod stegepanden, frem for en kokebog eller et tv. Man får med fx stemmekontrol mulighed for at bruge sine hænder og øjne til andre opgaver, end at betjene sig af en opmærksomhedskrævende, ekstern teknologi. Hørelsen er omnidirektionel, så brugeren behøver ikke at rette sin opmærksomhed mod et bestemt punkt for at tilegne sig informationer, men kan holde fokus rettet mod andre aktiviteter. Den lydbaserede kontrol gør brugeren i stand til at multitasker.⁹ Man kan sammenligne det med at lytte til et radioprogram, mens man bager pandekager, eller maler vinduer. Ved at anvende auditiv kontrol af systemer, er det muligt for brugeren at håndtere flere opgaver på samme tid.

Vi har taget udgangspunkt i VoiceXML-baserede browsere, der kan betjenes fra en telefon med håndfri funktion. Det er således ikke nødvendigt for brugeren at rette opmærksomheden direkte mod telefonen efter nummeret til Dinline er tastet, da al videre kommunikation foregår gennem den indbyggede højttaler og mikrofon.

Det er naturligvis en forudsætning at brugeren ikke er i tvivl om hvilke muligheder der er på det pågældende sted. Det nytter fx ikke at brugeren forventer, og forsøger at fremskaffe, informationer om holdenes placering i superligaen, når det nu er en tjeneste der guider brugeren gennem madlavning. Ligeledes skal der i de auditive kommandoer være en klar afgrænsning af applikationens funktionalitet. Det skal implicit fremgå af kommandoerne, hvad der er muligt.¹⁰

Derfor er det vigtigt at få præsenteret de mulige kommandoer ”pause, videre, gentag”, og gøre det med en anden speakerstemme end den der ellers giver informationer, så brugeren forstår, at det er brugeren selv, der skal sige det. Vi har desuden forsøgt at hjælpe udtalen lidt på vej, ved at lade den mandlige speaker udtale kommandoerne langsomt og med en lille pause mellem stavelserne.

⁹ Ronald Rosenfeld, Dan Olsen & Alex Rudnicky, *Universal Speech Interfaces*, Interactions, November + December, 2001, s.35

¹⁰ Ronald Rosenfeld, Dan Olsen & Alex Rudnicky, *Universal Speech Interfaces*, Interactions, November + December, 2001, s.38

I et afgrænset system som vores, giver dette simple sprog (kommandoerne: pause, videre, gentag, ja og nej) brugeren tilstrækkelig magt til at udtrykke alle nødvendige inputs.¹¹

Vores aktionssignal er et pift. Dette er valgt, for det første fordi et sådant pift også kendes fra hverdagen, når der er nogen der ønsker opmærksomhed. Piftet har altså en semantisk funktion der signalerer ”vær opmærksom/hør efter”. For det andet er det en lyd der adskiller sig markant fra de andre lyde i *Dineline*. Mindre hensigtsmæssigt havde det været, hvis vi fx havde valgt lyden af en klokke, da denne lyd kunne risikere at blive opfattet som lyden fra mikrobølgeovnen.

Endelig ligger der også en graduering i væsentligheden af informationerne som brugeren hører. Når det er kvindestemmen (den elektroniske guide) der *taler*, skal der høres godt efter, og når man hører en mandestemme finder man hurtigt ud af at informationerne er af mere sekundær karakter. Ved at anvende den samme speakerstemme til de væsentlige informationer, ved brugeren hvornår der skal høres efter.

For at gøre brugeren opmærksom på, at nu kommer der noget man skal tage stilling til, høres som nævnt et pift. Piftet er her en symbolsk lyd, da der ikke er en naturlig forbindelse mellem piftet og det piftet symboliserer, nemlig et forestående valg. Brugeren ved hvad piftet symboliserer, men piftet er ikke valgets lyd, da et *valg* er et abstrakt begreb, og som sådan ikke har en lyd.

På sin vis kan piftet dog også karakteriseres som værende en indeksal lyd, da piftet indikerer en sammenhæng mellem to fænomener. Det er ikke piftet der er det væsentlige, men den opmærksomhed, der kræves af brugeren i forbindelse med den forestående valgsituation.

Lyd er et flygtigt fænomen. Den kan ikke høres, når den er holdt op. Dette har naturligvis betydning for designet af det interface som skal betjenes ved hjælp af lyd. Lydens flygtige karakter betyder, at der er grænser for hvor mange niveauer en bruger kan forventes at kunne huske, da der, som i vores tilfælde, ikke er en stifinder, der kan minde brugeren om hvor han er, og hvordan han kom dertil. Det er en af de væsentlige årsager til, at det er vigtigt at skabe redundans, så brugeren ikke er i tvivl om hvor han er på vej hen. Dette gør vi ved hele tiden at gøre opmærksom på, hvilke valg brugeren har truffet. Fx ”Du har valgt at høre en scorehistorie”.

De fleste af os kender de windowslyde der optræder når vi for eksempel foretager en ’ulovlig handling’, eller tømmer ’papirkurven’. Det er denne type lyde som kaldes for bekræftende lyde¹². Disse lyde har til

¹¹ Ibid. s.39

hensigt at gøre det lettere for brugeren at gennemskue hvilke handlinger han udfører. Vi benytter os ikke af bekræftende lyde, da navigationskompleksiteten ikke er alt for omfattende. Når man i vores voice-browser giver kommandoen *videre*, fungerer selve fortsættelsen som den bekræftende lyd. Det ville derfor være overflødigt at tilføje yderligere lyde der skal bekræfte brugeren i den udførte handling.

Sprog

Hovedelementet i vores produkt er stemmen, systemets ”stemme” såvel som brugerens stemme.

”Stemme” oversættes som bekendt til ”voice” på engelsk. Slår man ”voice” op i Gyldendals engelsk-dansk ordbog (og det gør man, underligt nok, selv om man jo lige selv har oversat det fra dansk), står der

der	skrevet	nogenlunde,	som	følger:
-----	---------	-------------	-----	---------

Voice; sb stemme, røst; vb. give udtryk for, (gram, om verber) genus, form, diatese; *the active* ~ aktiv, handleform, *the passive* ~ passiv, lideform, (fon) stemme, udtale, *-d* stemt.

Siger man derfor ”voice”, som - for nu bare at tage et helt tilfældigt eksempel - i ”Voice XML”, menes der altså ikke nødvendigvis bare ”stemme”, som i ”sikken en flot stemme Aretha Franklin har”, men også noget om grammatiske former og fonetik.

I udviklingen af ”Dineline” har vi i høj grad måtte tage stilling til stemmen som andet og mere end bare et substantiv. (For nærmere gennemgang af de fonetiske overvejelser, se afsnittet om VXML). Først og fremmest har vi selvfølgelig taget stilling til, *hvad* ”stemmen” skulle sige (se bilag for manuskript), dvs. semantikken. Vi skulle på en eller anden måde have brygget en instruktion sammen, som gav mening, for den der skulle bruge systemet.

Vores system er et interaktivt system, og brugeren skal således interagere med systemet. Måden, man interagerer med dette system på, er ved hjælp af stemmen. Interaktion med stemmen minder til forveksling om en samtale. Vores produkt efterligner da også i høj grad samtalen, men *er* ikke en samtale, da den ene af parterne er elektronisk. At efterligne samtalen er oplagt og letter til en vis grad, gennem sin genkendelighed, interaktionen, men stiller også i høj grad både udviklere og brugere overfor nogle udfordringer.

Allerførst var vi nødt til at tage stilling til, hvilken bruger det var, vi stod overfor (se afsnittet om målgruppe). Både for at finde ud af, hvilket sprog vi ville gøre brug af (og her tænkes ikke på engelsk, tysk eller fransk, men ”knæhøj karse”, ”totalt for svedigt” eller ”helt tip top”), samt hvor meget brugeren vidste og kunne finde ud af. F.eks. valgte vi i demonstrationsmodellen at spørge brugeren, om

¹² William W. Graver & Randall B. Smith, *Human-Computer Interaction*, Interact '90, s.565

han ved, hvad et vandbad er, men spørger ikke om han ved, hvad et ildfast fad er. At blive spurgt om man ved, hvad noget, som man selv opfatter som meget enkelt, er, kan meget let virke lidt irriterende. Går folk derimod ud fra, at man udmærket godt ved, hvad der tales om, når man faktisk ikke gør det, gør det ikke bare én irriteret men også én en lille smule hjælpeløs. Ved man her i vores system godt, hvad et vandbad er, kan man bare blive lidt irriteret og svare ”jo”, når man bliver spurgt, så går systemet videre med instruktionen. Ved man til gengæld overhovedet ikke, hvad et ildfast fad er, er der ikke megen hjælp at hente fra systemet. (Så må man ringe til sin mor, på nær at man jo bruger telefonen...). Et andet eksempel, på at vi tager stilling til, hvor meget brugeren kan, er, hvornår vi vælger at indlægge pauser i instruktionen. F.eks. siger vi alt om tilberedningen af salaten i én køre, fordi vi opfatter det som noget, ”man sagtens kan finde ud af”, hvorimod vi venter på en ”videre”-kommando fra brugeren, efter han har fået at vide, at han skal ”skære knoldene af champignonerne”. Modsat eksemplet med det ildfaste fad, kan brugeren dog her gøre noget ved situationen ved enten at sige ”pause”, ”videre” eller ”gentag”. At det ikke *er* en almindelig samtale, ses selvfølgelig på de strikte kommandoer brugeren er tvunget til at sige (han kan f.eks. ikke vælge at sige ”vent lige lidt” i stedet for ”pause”), men afspejles der ud over også i, at systemet ikke lader sig påvirke af brugerens humør. F.eks. siger brugeren irriteret ”videre for fanden” efter instruktionen: ”Hold med højre tommel og pege på stilken, og dyp jordbærrene halvt ned i den smeltede chokolade.”, hvortil systemet kækt og upåvirket svarer: ”Sæt jordbærrene på dit *fineste* fad og stil dem i køleskabet”. Ydermere er turn taking¹³ heller ikke helt naturlig. Godt nok kan man som bruger til hver en tid afbryde systemets talestrøm, men systemet har ikke kontakt med/ fornemmelse for bruger (fx via øjenkontakt), pauserne er planlagt på forhånd, og bruger kan ikke lede systemet ”ud af en tangent”.

For at undgå fejl, misforståelser og usikkerhed ville vi gerne tydeliggøre, at brugeren kun har et begrænset antal kommando-muligheder. Vi startede således, i introduktionen til systemet, med at sige, at ”når denne lyd kommer [piftet] , vil du blive stillet overfor nogle forskellige valgmuligheder, hvortil du skal svare ja eller nej”. Dette måtte vi senere ændre til: ”hvortil du skal indtale dit svar”. Grunden hertil var, at svar måske kunne opfattes som generelle kommandoer (”videre”, ”pause” og ”gentag”), da det måske kunne ses som ens ”svar” til systemet. Ydermere spørger vi på et tidspunkt: ”Ved du slet ikke, hvad et vandbad er?”, og her ville det være naturligt at svare ”jo” i stedet for ”ja”, hvis man altså ved, hvad det er. Man kunne også have valgt at formulere spørgsmålet som: ”Ved du, hvad et vandbad er?”, så brugeren hertil kunne svare enten ”ja” eller ”nej”. Generelt gjorde vi os en del overvejelser omkring lige netop den sætning. Sætningen optræder midt inde i instruktionen, og under optagelsen

¹³ Turn taking er et begreb fra Speech Act Theory (John Lakoff) og henviser til de naturlige ”snakke-ture” man skiftevis tager i en samtale.

blev det tydeligt, hvor stor betydning intonation har for interrogativer. Glemmer man at ændre sit toneleje ved en interrogativ, er det næsten umuligt for modtager at opfatte meningen. Og selvom vi havde lagt et **pift!** ind umiddelbart før sætningen, blev vi nødt til at tage optagelsen om og lave lidt om på ordvalget, så det passede bedre ind i stilen og blev nemmere for speaker at formulere som et spørgsmål.

Det meste af scenariet består af korte, imperative sætninger, altså en slags ordre. Dette er helt naturligt, når man tænker på, at det er en instruktion, men understreger også det faktum, at det ikke er en samtale. F.eks. har vi snakket om, om vi skulle have en tilbage-funktion/ kommando. - Som systemet er i øjeblikket, kan man nemlig komme ud i problemer, hvis man bare lader instruktionen køre som en slags baggrundsstøj, mens man fx snitter peberfrugt, for man kan ikke gå mere end en sekvens tilbage. Vi ligger altså op til, at man hele tiden skal være vågen, - og et evt. fravær kræver interaktion med systemet først, nemlig kommandoen ”pause”.

En anden kommando vi har diskuteret lidt, er kommandoen ”videre”. ”Videre” vil af systemet, som det er i øjeblikket, altid blive opfattet som ”videre med opskriften”. Siger brugeren derfor f.eks. ”videre” under scorehistorie nummer et, fortsætter systemet ikke med scorehistorie nummer to, men går videre med opskriften. Man skulle måske gøre opmærksom på denne feature i introduktionen til systemet, hvis man ville gå videre med produktet.

Konstruktionen

Vi har lavet en længere produktion, den er sammenlagt på mere end 12 minutter. Vi har lavet al speak selv, hvilket tog en hel del tid. Ikke mindst fordi vi optog nogle ting flere gange, for at have noget at klippe imellem. Vi havde i ny og næ problemer med indstillinger af det lydudstyr, vi havde til brug. Hver gang vi optog forskellige personer, tog det lidt tid, før vi fandt en indstilling, der passede til personen. Det endelig mix Cool Edit Pro lavede var heller ikke helt brugbart, men skulle normaliseres et par gange, for at undgå at der kom for meget overstyring på enkelte dele. Når Cool Edit Pro laver et ”mixdown”, laver den et gennemsnit, der gør at det bliver tilladt for enkelte lyde at gå over grænsen. I vores produktion har det typisk været enkelte ord der har skilt sig ud. Enkelte steder er det fordi vi ikke under indspilningen har været opmærksomme på, at en stavelse har været for høj (den har peaket) som input. Nogle gange har vi kunnet redde det ved at normalisere ned, andre gange har vi været magtesløse.

Som lydeksperter Eddy Bøgh Brixen, der til hverdag arbejder på ”KKDK acoustic”, gjorde opmærksom på under et foredrag om lydproduktioner d. 4/3 i forbindelse med kurset, er det bedst at

konstruere sin produktion i højeste kvalitet, og så senere gå på kompromis med ”virkeligheden”. Det er et over-kill af dimensioner at bruge 44 kHz, al den stund telefoner slet ikke sender i den kvalitet. Men da vi kun har lavet et scenarie for brug, ikke en endelig løsning, syntes vi godt, at vi kunne forsvare det. Vores produktion fylder samlet 125 MB, hvilket er en hel del. Men vores produktion skal ikke streames over Internet, men over telefon. Her er det ikke størrelsen, men gørelsen der er af betydning. Størrelsen har betydning for den server, der skal afvikle programmet, men man kan ”spare” mange bits ved at bruge korte loops. Det har vi gjort undervejs, men Cool Edit Pro (som vi har benyttet) tager ikke hensyn til, om en lyd afspilles en eller flere gange. Et loop fylder for hver gang, det bliver brugt i et mix.

Vores lydproduktion er ikke real-time, men det er tæt på. Vi har valgt at lave en illusion om, at brugeren rent faktisk laver de enkelte dele, ved at lave meget korte muzak-loops og så indsætte ”videre” for at få systemet startet igen.

Vores lydproduktion består af en lang wav-fil. Den er resultatet af sammenklipning af et utal af småloops og bidder.

Alene indgangen til programmet består af tre bidder – en stemme der siger ”velkommen til”, en lydfil indspillet på midi-keyboard og vores lydlogo.

Der har været problemer med enkelte af de lyde, vi har inkorporeret. Brugen af hajens ledemotiv i ”Jaws” er animeret af billedet som svigermor som noget man frygter. Ideen er meget god, men det eneste klip med Jaws, vi kunne finde var i en ret dårlig udgave, 44 kHz/ 14H kbps, med en hel del hvid støj, hvilket gjorde det lidt svært at arbejde med. Hvis det blev normaliseret ”for meget” op, kom der for meget støj med. Vi er havnet på et kompromis, hvor ”svigermor kommer” i sig selv er blevet delvist normaliseret. Der er her tale om et psykoakustisk fænomen, da det lyder lidt lavere end de andre lydbidder i menuvalget, på trods af at det er relativt højt volumenmæssigt (målt i DB, er det lydclip det højeste sted i vores produktion). Som det er i vores produktion lyder det lavere, fordi baggrundslyden er uskarp. Vi kunne måske være gået lidt uden om problemet ved at lægge jaws-temaet ind to gange, panorere det ud og så have selve talen centreret. Men det væsentlige i dette tilfælde er ikke udførelsen, men ideen.

Hvis det havde drejet sig om instrumenter, havde man nok valgt at bruge flere tracks for at have muligheden for at bruge en generel panorering eller volume-setting til det enkelte spor. Men for at gøre det lettere for os at se, hvor de enkelte spor lå i det temporære forløb, har vi kun brugt fire. De er

temporære lydligt, men vises for os spatialt i programmet. Men her er det kun et spørgsmål om grafisk repræsentation.

Vi har lagt nogle session-filer ind i selve produktionen, men de var ret korte tidsmæssigt og let overskuelige. De kan relativt smertefrit genfindes, da Cool Edit Pro har gemt dem som filnavn-Mixdown.wav.

Grunden til at vi har haft et liberalt forhold til placeringen af de forskellige lyde er, at brugeren jo aldrig vil komme til at se sekvenserne som grafik. Det eneste brugeren interagerer med er lyd, så inde i Cool Edit Pro er informationsarkitekturen ikke så vigtig.

Ved at lægge session-filer ind i dineline.ses., afskar vi os fra at redigere i mindre enheder, og generelt er det ikke så praktisk at bruge mixdowns af ”sessions” inde i nye filer, men det har gjort det lettere at arbejde med.

Vi har indspillet de ord, som brugeren siger, et par gange, og så klippet dem lidt tilfældigt ind for variationens skyld. De spiller ikke den store rolle i produktionen, men skal mere illustrere, hvordan systemet virker. Brugers svar skal vise, hvordan de forskellige funktioner virker. Vi har kun brugt ”gentag” en enkelt gang, men denne funktion ændrer sig heller ikke yderligere. Vi ville bare vise, at systemet forstår tale. Lyde som ”ja”, ”nej”, ”videre”, der bruges mange gange i forløbet, har vi indspillet i flere versioner, udelukkende for at gøre scenariet mere interessant at høre på.

Den lange ”samtale” mellem brugeren og speakeren er lagt ind som en lang wav-fil. Det har vi gjort fordi, der ellers skulle flyttes rundt på mange forskellige dele, når der blev indsat nye, uventede elementer. Det betyder ikke andet, end at alt det, der høres som samtale, er klippet ind til et spor. Filen blev godt nok så lang, at Cool Edit Pro har haft visse problemer med at håndtere den, men de problemer hører til i småtingsafdelingen.

Cool Edit havde også nogle problemer med at læse enkelte filer, den selv havde konstrueret. Hvis man åbner Dinline.ses, skal man acceptere at enkelte filer læses på en bestemt måde. Det er sket fordi, nogle filer på et tidspunkt er blevet gemt som pcm-wav-filer og ikke windows-wav. Denne fejl er ikke blevet rettet, og efter at de så er blevet lagt ind i en session, kunne de ikke efterfølgende ændres.

Vi overvejede at bruge reallyde, men gik væk fra det igen, da det ville gøre vores produktion mere dramatisk. Det ville også gøre, at vores scenarie blev fokuseret på begge sider af røret. I første omgang er vi interesserede i, hvad brugeren hører, ikke hvad der kom tilbage til systemet.. Derfor har vi lagt brugers stemme ind, men vi har ikke forsøgt at imitere virkelighedens verden ved at lægge lyde fra et køkken ind.

Dineline – musikbrug

Anvendelse og omfang

Der er brugt musik i hele produktionen Dineline. Den fungerer på den måde, at der i standardindledningen optræder et lydlogo og en række vignetter, og der i selve programmet indgår decideret underlægningsmusik.

I programmet skelnes der mellem to forløb; det der er præindspillet, såsom ”drik en øl”, ”mere om vine” og ”scorehistorie”, og text-to speech forløb hvor den elektroniske guide instruerer. Her forestiller vi os, at hver kommando-prompt eller trigger har tilknyttet et musikforløb eller loop. Når brugeren giver en kommando fades dette afsnit ned, hvorefter næste episode starter.

Vores prototype bærer i nogle henseender præg af, at vi er gået på kompromis mellem at fremstille et troværdigt forløb, og et velklingende produkt. Når brugeren giver en kommando i programmet, lader vi for eksempel musikken fade vilkårligt, stort set uden overvejelser om æstetik og dramaturgi. Til gengæld genbruger vi ofte numrene for ikke at gøre produktet for forvirrende. I et realistisk tidsforløb ville brugeren nok foretrække, at baggrundsmusikken skiftede oftere.

Musikkens funktion

Musikken i programmet har en række praktiske og æstetiske funktioner i programmet.

Hensigten med overhovedet at inddrage musik er helt enkelt at lette forløbet for brugeren, som i vores tilfælde vælger billig, begynder og derfor næppe er den store entusiast. Først og fremmest har vi derfor lagt vægt på, at det skal være sjovt at lave mad, og har derfor efterstræbt at programmet skal gøre madlavningen underholdende og nemmere.

Programmets underholdningsværdi eksisterer blandt andet i kraft af den overordnede stil, som kort fortalt er en blanding af mild satire, TV-køkken- og spøjs retrostemning. Programmets stil og stemning dikteres i høj grad af musikken, og denne er derved en gennemgående stilsætter for programmet.

Musikken kan derfor siges overordnet at være parafraserende, idet den understøtter den handling, der er i programmet, og den overordnede stilmæssige idé der ligger til grund herfor. Oftest vil musikken dog opleves som polariserende. For eksempel kan dens brug i satirisk øjemed måske virke direkte provokerende (fx Poul Dissings kommentar ”Amore” i scorehistorien), men helt generelt sker det i og med at den benyttes meget bevidst så brugeren påvirkes til at opleve noget specifikt. På denne måde har musikken også en funktion som pædagogisk og dramaturgisk værktøj.

Det er endvidere hensigten, at musikken skal fungere som kapitelinddeler eller episodemarkør i programmet og skal kunne hjælpe brugeren til at vide, hvornår der sker noget nyt, og hvor vi er i ”det

auditive rum” – i køkkenet ved madlavningen, eller om man befinder sig midlertidigt ”udenfor” i kraft af en scorehistorie, pause eller andet.

Valg af musik

Underlægningsmusikken er generelt udvalgt for sine lydlige og musikalske kvaliteter i forhold til at kunne etablere en stemning. Gennemgående har musikken nogle meget stærke stilindikatorer eller træk, som i forhold til sin genre er meget stereotype. Mange af de musikalske elementer fremstår derved som klichéer, hvilket understøtter dels ironi, satire og retroelementet, dels en fornemmelse af muzak og baggrundsmusik og dermed de situationer, hvor dette indgår. Vigtigt for det auditive udsagn af programmet er også tempoet i musikken, tonehøjder og andre elementer, som relaterer til noget fysisk hos brugeren, og som man ubevidst reagerer på. For eksempel at noget musik giver én lyst til at danse, andet får én til at slappe af osv. Musik og lyd kan desuden give en illusion af en bestemt rumlighed.

Derudover er der lagt vægt på, at musikken enten er ukendt i forhold til målgruppen, eller i sin redigerede form fremstår nogenlunde anonymt, således at brugerens kendskab til, eller oplevelse med, det konkrete musikstykke fra andre sammenhænge ikke er forstyrrende for musikkens betydning i dette program. Modsat er det med lydlogoet og vignetterne i standardindledningen, hvor netop brugerens referenceramme til musikken udnyttes.

I det følgende vil anvendelsen af musik blive gennemgået med særligt henblik på musikkens betydning for de enkelte afsnit, samt hvordan dette står i forhold til programmet som helhed.

Lydlogo

Vi har lavet et lydlogo der meget bevidst læner sig op ad vignetten til programmet ”TV-køkkenet”. Det er en parafrase over temaet, men alle der har hørt det, har ”genkendt” det. Vores tilføjelse består af en ny slutning og nogle simple vokalharmonier. Et firstemmigt kor synger akkorderne G og A på teksten ”Dineline”. Fire stemmer synger altså tilsammen tre toner ad gangen, hvoraf 2 stemmer stiger og to falder, og vi har derfor panoreret sporene ud til hver side for at få koret til at fremstå mere klart og få en fyldig effekt. Rent teknisk er sporene heller ikke helt synkrone, da de er indspillet uden medhør, hvilket gør det ret svært at lave mere end en omtrentlig indspilning. Hvis man spiller de fire spor centreret, kan man tydeligt høre, at de ikke er helt lige lange, men når de er panorerede, bliver forskellen mindre.

Vores lydlogo er relativt langt, næsten ti sekunder, hvilket er alt for meget. Den del hvor navnet nævnes er dog kun på to sekunder, så hvis navnet bliver mere kendt, kan man nøjes med det, ligesom andre

firmaer (eksempelvis Mc Donalds), der har fået cementeret sit lydlogo via intens brug af reklamer. Det ses ofte i reklamebranchen, at man i begyndelsen af en kampagne bruger en lang blok, og senere skifter til en kortere variation, når først forbrugerne er vænnet til reklamen. Eksempelvis kører firmaet TDC med to udgaver af en reklame, en lang og en kort. Man ser det også ofte i TV-serier ved scenskift (som fx i "Friends").

Grunden til at lydlogoet er så tæt beslægtet med "TV-køkkenet" er, at vi ville have brugeren til at tænke på mad fra første øjeblik. Selvom bemeldte program ikke har kørt i en hel del år, er det et lydlogo de fleste over en hvis alder kan genkende. Man kunne argumentere for, at folk under 16 måske ikke kan huske det, hvilket med tiden ville give nogle identifikationsproblemer.

Førstestemmen i vores lydlogo består kun af to toner, og de andre stemmer er lavet i harmonier, der næppe kan støde nogen. Stilistisk ligger det meget tæt på reklamespots for amerikanske firmaer. Man kan argumentere for at der er så meget vellyd, at det kan være svært at huske, men så kunne man til gengæld lave lidt disharmoni i et baggrundsspor. Et trick der bruges meget i popmusik for at tiltrække opmærksomhed på bestemte tidspunkter. En kunstner som Alanis Morissette har brugt dette kneb en del gange, blandt andet i gennembrudshittet "You oughta know".

Nogle brugertests af vores produktion viste at testpersonerne havde problemer med at høre de ord, der bliver sunget. Men det handler nok i nogen grad om forankring. De folk, der har hørt vores scenarie, har ikke på forhånd vidst, hvad de gik ind til, mens man må formode, at folk der ringer ind til tjenesten ved, hvad det er, de ringer til.

Vignetter

I standardindledningen præsenteres en række menuer, som brugeren skal vælge imellem. For at skabe noget suspense i programmet, har vi valgt at præsentere disse menuer som vignetter.

Vignetterne skal understrege titlen på den enkelte menu, og i det hele taget fungere som appetitvækkere til netop deres programindhold. Vignetterne er samtidig, udover lydlogoet, det første musik, man hører i programmet, og det er derfor vigtigt, at deres udsigelse også stemmer overens med hele programmets stemning.

Vi har derfor valgt at lade parodien være stærkt repræsenteret her, hvilket dels kommer til udtryk gennem stemmeføringen, dels gennem valget af musik og dels gennem klipningen.

For eksempel siges titlen ”svigermor kommer” med en masse tremolo på stemmen, musikken til ”madpakke” er overdrevent gammeldags og barnlig i forhold til målgruppen og i ”scoremiddag” er musikken klippet til, så kun det første riff fra sangen ”Kiss” – en sang de fleste kan referere til - er med.

Dagens menu (jazz)

Efter standardindledningen og en bekræftelse af brugerens valg, følger en præsentation af dagens menu, og det er her, det egentlige program starter.

Som underlægning har vi valgt at bruge et 4-takters tromme loop i hurtig swing-stil som kører ca. to gange, med mindre brugeren interagerer.

I hvert af loopets takter er der en klar optakt, og i den fjerde takt markeres yderligere slagene 1 – 2 – 3 – 4. Derudover er instrumentationen i loopet udelukkende whiskers, og man forventer derfor, at der vil komme noget mere. Idet loopet i stedet gentages, samt det hurtige tempo taget i betragtning, bærer hele musikforløbet derfor alt i alt præg af noget utålmodigt eller uforløst.

Loopet skal derved fungere som en igangsætter til programmet, som bereder brugeren på at komme op i omdrejninger, men samtidig skal det også indikere, at vi stadig er ved en præsentation og ikke selve madlavningen.

Den bratte afslutning har til hensigt at skærpe brugerens opmærksomhed overfor det følgende programpunkt.

Ingredienser

Efter præsentationen af dagens menu indgår en sekvens, hvor ingredienserne gennemgås. Vi forestiller os, at denne del er præindspillet.

Underlægningsmusikkens funktion her er først og fremmest at tilføre programmet et underholdende element, og skal ligesom vignetterne i indledningen fastslå den satiriske stil fra starten.

Sekvensen er en stiliseret udgave af en indkøbssituation.

Dette opnås blandt andet gennem musikken, hvor vi har valgt et stykke underholdningsmusik, som indeholder en masse musikalske referencer til den muzak, der spilles i supermarkeder. For at parodiere noget husmoderligt fremlægges ingredienserne som en remse henover musikken. Dette foregår i en rytmisk, stacceret oplæsning, som er kørt igennem en equalizer for at opnå en lyd af gammeldags radio. Det vokale udsagn understøttes yderligere af musikkens mange episodemarkører og optakter, som hele tiden opdeler forløbet og bringer det videre. Sammen med musikkens puls, et raskt gå-tempo, og den firkantede opdeling i lige taktantal og stykker, får hele sekvensen et præg af effektivitet og systematik, som det endvidere er tiltænkt skal smitte af på brugeren.

Pause

Ved øl-pausen vendes tilbage til jazzen, dels for at afslutte denne indledende del af programmet, men også for igen at placere brugeren i programmet efter evt. indkøb og at få brugeren frem i skoene. Desuden sammenkædes jazzen ofte med hygge og med baggrundsmusik, og er derfor også valgt med henblik på at sætte brugeren i en afslappet stemning.

Pandekager (Bagdad Dagblad)

Nu starter programmets hovedindhold, nemlig instruktionen. Musikken er her tilknyttet bestemte instruktioner i programmet, og optræder derfor på baggrund af brugerens kommandoer. Musikkens dramaturgiske egenskaber er derfor nedprioriteret i forhold til dens pædagogiske og afsnitsinddelende evner.

Musikken indledes med en fanfareagtig og pludselig optakt for at indikere, at nu skal der bestilles noget. Herefter fortsætter musikken i et tempo, som er moderat, men som synes pågående på grund af den stærke markering af alle taktslagene. Igen er dette hensigtsmæssigt med henblik på at holde brugeren i gang. Temaer og melodi i musikstykket består endvidere af få, rytmisk distinkte stykker, hvilket virker godt i forhold til at illudere systematikken i en madlavningssituation.

Musikkens stil er en form for fusionsjazz og indeholder forskellige referencer til musikalske stilarter, men også til mere specifikke ting. For eksempel er en del af temaet instrumenteret med elguitar og vibrafon, hvilket er brugt meget i blandt andet vignetmusik og populærmusik fra 60'erne og 70'erne, og derved giver dette musikstykke også et præg af parodiering og en retrostemning.

Samtaleråd

Musikken til "Samtaleråd" er underlægning til en fortæller, som gerne skulle fremstå smart, tjekket og troværdig. Vi har valgt at benytte et loop, som er og virker mere moderne, og dets tilbagelænedede beat skal være med til at opbygge en stemning hos brugeren om at have tjek på tingene. På den måde understøtter musikken samtalerådenes funktion helt.

Musikken optræder i dette afsnit væsentlig højere end underlægningen til instruktionerne, mens indtalingen er lavere og med større rum i forhold til den elektroniske guide. Lydens niveauer er på denne måde væsentlig for at markere dette nye afsnit, og sammen med musikkens tjekkede stil udnyttes det yderligere til at placere brugeren i et nyt "rum" i forhold til det retroagtige TV-køkken.

Fyld (Henri Mancini)

Vi er nu tilbage i programmet, og madlavningssituationen understreges ved, at musikken vender tilbage til jazz-underholdnings-muzak stemningen. Idet brugeren nu må forventes at være godt i gang med kokkereringen, er det pågående og systematiske i musikken her nedtonet, men tempoet er dog stadig oppe.

Scorehistorie

Som musik til scorehistorien har vi valgt tango, som med sin vekslen mellem blid legato og voldsom staccato er indbegrebet af stor, dramatisk musik.

Som det var tilfældet med samtalerådene, er lydniveauer på indtaling og musik samt musikkens fremmede stil med til at adskille afsnittet fra instruktionerne. Ellers er musikkens primære funktion her at skabe suspense og drama.

Historien indledes med noget nært en musikalsk kliché, en tangofigur, som her fungerer som optakt til historien og skal indstille brugeren på at høre en dramatisk historie. I sin oprindelighed indgår denne figur en række gange som eftersætning til forskellige passagerer i tangoen, så idet vi indleder med dette, kan man sige at historiens gang foregribes.

Hele vores brug af musikken i dette afsnit er stærkt overdreven, og også ironisk tiltænkt - især hen imod slutningen, hvor fortællingen er standset, men musikken fortsætter og accelererer i tempo. Ved denne ironiske musikalske kommentar til historien bevæger musikken sig over til næsten at stå i et kontrapunktisk forhold til talen.

Fyld2 (Henri Mancini)

I denne del af instruktionen er det forholdsvis nemme, og få, ting der skal gøres. Her er musikken igen den parodierende muzakstemning, men denne gang noget roligere og lidt selvtilfreds, så brugeren kan slappe af.

Pandekagerulning (Eels)

At rulle pandekager med fyld er svært og kræver fuld opmærksomhed. Vi har, for at holde stilen, valgt at stresser brugeren ved at bruge et loop, som er pågående og uroligt – tempoet er hurtigt, de enkelte stykker virker mystiske med riffs, som rykker skiftevis en halv tone op og ned, høje orgeltoner, dyb bas og skarpe panoreringer.

Jordbær (klassisk)

Her skal man igen gøre sig umage, men på den sofistikerede måde, og vi har brugt klassisk musik, opera i en let stemning. Desuden er der her tale om menuens absolut mest romantiske indslag, hvilket gerne må underbygges af musikken.

I forhold til de øvrige instruktioner, som har været baseret på musik, som har været rytmisk funderet og har haft mange fællesnævner i instrumentering og form, sker der her et musikalsk stilskifte.

Systematikken i afsnitsinddelingen bliver dermed brudt, men dette har til dels et pædagogisk formål idet det markerer, at hovedretten er færdig, og næste ret kan påbegyndes.

I sin sammenhæng virker musikken parafraserende, men overordnet er den i kraft af stilskiftet polariserende. Idet musikken fungerer som underlægningsmusik falder den dog fint i spænd med muzak-referencerne. Den malplacerede stil i forhold til det øvrige musik og i forhold til målgruppen og den overdrevne romantik, gør yderligere at ironien bibeholdes, hvorfor der ikke er tale om et egentlig stilbrud.

Salat (Beck)

Musikkens stil er her en mere moderne og tilbagelænet stil, som i samtalerådene, og der er i forhold til de øvrige instruktioner derfor igen tale om et stilskifte.

Stilskiftet skal indikere at endnu et menupunkt, salaten, påbegyndes.

Musikken er desuden udvalgt for på falderebet at imødekomme målgruppens musiksmag, og for at genopbygge den selvtillidsopbyggende stemning fra samtalerådene.

Færdig (fanfare)

Når brugeren er nået igennem alle instruktioner, som har at gøre med madlavningen, lyder en fanfare.

Dette er valgt, fordi vi går ud fra, at det har været en overvindelse for brugeren at kreere en middag, og vi vil derfor gratulere.

Fanfaren forløber over 18 takter, hvor temaet varieres og bygger op til klimaks, som i forhold til vores handling er: Er brugeren klar, ja eller nej?

Fanfaren er igen et stilskifte i programmet, starter endvidere meget bombastisk og er i det hele taget overproportioneret i forhold til den øvrige musik. Formålet hermed er klart at markere afslutningen på madlavningsdelen, men også for slutteligt at understrege ironien med fed.

Slut (musik + logo)

I afslutningen forestiller vi os at programmet spiller en række musiknumre som baggrundsmusik mens brugeren gør sig klar. Når brugeren siger ”videre” afsluttes programmet med lydlogoet som sidste indslag.

Ophavsret

Ved at bruge de samme loops hver gang, er det muligt at udregne meget præcist, hvor meget hver sang skal have i KODA-penge. I princippet er det samme afregningsmetode, som radioer bruger, med en spilleliste med angivelser af antal spillede minutter. Det er principielt ikke svært at lave en meget præcis afregning, da al lyd bliver spillet fra en server, så man let kan betale per brugt enhed/loop.

Eftersom vi benytter os af få kunstnere, ville det letteste være at lave en fast aftale. KODA har en liste med priser til forskellig brug, og her fremgår det at ventesignaler på telefon til 20 linjer koster 1.146,00 kr. til KODA og 745,00 kr. til Gramex årligt.¹⁴ Der er ingen præcedens på området, da der ikke eksisterer et system, der ligner vores, men man kunne forestille sig, at KODA ville gøre brug af en anden form for afregning, eksempelvis en der ligner den man bruger i forretninger, der betaler en mindre afgift hver dag, for at spille musik fra radio eller muzak.

Afrunding

Vores udgangspunkt var, at vi ville lave et koncept, der var baseret på nogle af de muligheder, som VoiceXML byder på. Princippet er som tidligere nævnt ikke nyt, men sætter man det sammen med talesyntese og text-to-speech, åbner der sig nogle muligheder som vi har forsøgt at sætte i spil igennem udviklingen af vores koncept.

Der har altså fra begyndelsen været tale om et koncept, der er mere eller mindre futuristisk.

Da teknologiens muligheder og perspektiver er mange, er der følgelig tilsvarende mange vinkler vi kunne have lagt på vores koncept, og mange andre elementer man kunne have inddraget og overvejet.

Vi kom for eksempel lidt for sent i forløbet på den idé, at det faktisk er muligt, ganske gratis, at anvende funktionsdygtige talesynteser, der i stemmekvalitet ligger langt fra den typiske robotforestilling om talesyntese. Det ville have været oplagt at anvende en sådan talesyntese i vores produktion, for derved for os selv at have fået en klarere fornemmelse af hvordan sådan en lyder og fungerer, og for

¹⁴ <http://www.koda.dk/usr/koda/kodaweb.nsf/index>

det andet havde vores indspillede scenario/mock-up virket mere realistisk. På den anden side, er det ikke umuligt at forestille sig, at når en sådan service, som vores DineLine kunne blive tilgængelig, vil der ikke være en umiddelbar hørbar forskel på talesynteser og ægte tale.

Tilsvarende virker det heller ikke urealistisk at forestille sig, at der inden for en overskuelig fremtid kommer mobile (og fedttætte) fladskærme, som kunne fungere sammen med DineLine, hvorpå der så kunne være en skreven udgave af opskriften, ingredienserne, billeder af hvordan retten skal se ud og måske billeder af hvordan man gør.

Havde vi haft flere ressourcer (tid og penge) og assistance fra kyndige programmører, ville det ikke have været urealistisk at lave en egentlig prototype. Med en sådan havde vi først virkelig kunnet afprøve vores koncept, og fået mulighed for at lave egentlige usability tests. En usability test ville givetvis kunne bidrage med væsentlige erfaringer om lydbaseret HCI, da det ville være muligt at teste systemets egnethed i realistiske omgivelser.

Angående udformningen af vores konkrete produkt og specificeringen af konceptet er der ligeledes utallige andre veje man kunne have valgt.

Man kunne for eksempel forestille sig at DineLine i endnu højere grad var en del af et multimedie-koncept, hvor det samme produkt blev brugt på forskellige platforme såsom TV, radio, Internet og teletjeneste. I så fald ville det være yderligere relevant at tilpasse brug af musik og indslag til en målgruppe, og man kunne sagtens forestille sig at programmet præsenterede kendte mennesker, aktuel musik og rigtige interviews.