

Stilhed før fremtiden

*Af Lars Kai Hansen,
Danmarks Tekniske Universitet*

Min farfar yndede at fortælle en historie, om første gang han hørte lyden af en flyver på himlen i sin barndomsby Horsens. En uventet lyd fra fremtiden, der fik drengene til at stimle sammen og kikke på hinanden: Nu flyver vi! En dag i august 2004 besøgte jeg en avishjemmeside og hørte for første gang lyden af internettets fremtid. Da pilen passerede en reklame begyndte lyden pludselig at strømme ud af computeren. Uventet og varslende: Nu larmer vi!

Som andre førstegangsoplevelser på nettet, vil jeg huske denne lyd. Som min farfar vil jeg fortælle mine fremtidige børnebørn om denne augustdag på nettet: »I tror det ikke - men engang var der helt fuldkommen stille på nettet...«

Det er faktisk påfaldende så lidt vi bruger lyd på nettet i dag. Hvor der ellers er en nærmest ubændig vilje og trang til kommunikation. Vi bliver tæppebombet med reklamer, spam-mails, og pop-up's. Kampen om opmærksomhed er ubarmhjertig. Men endnu har kun ganske få opdaget, eller måske snarere, turdet, bruge lydens evne til at fange opmærksomhed ved computeren. Reklamerne på tv er avancerede mediemanipulationer, der arbejder med vores ubevidste trang til at forstå og huske lyd og billeder. Det er som om at vi har accepteret aggressiv brug af lyd på tv, men stadig ikke helt tilsvarende er parate til at lægge øre til internettet. Forklaringen kunne måske findes i den traditionelle brug af pc'en, hvor man sidder i stuen eller i et kontormiljø med flere personer.

På kontoret ville det være pinligt, og helt sikkert tegn på en svag karakter og mangel på koncentration, hvis computeren pludselig plaprede løs med reklamer. Men nu er computeren jo ikke længere bare et arbejdsredskab. Den er hurtigt blevet en personlig mediemaskine: musikanlæg, dvd-biograf og telefon. pc'en er vært for en masse funktioner med lyd på, og mange sidder i dag ved computeren med et headset. Så lydturen falder snart. Internettet kan og vil eksplodere i kakofoni.

Internettet har i dag to sider. Oversiden som man umiddelbart møder, når man surfer rundt på nettet. Og så den anden side, den del af nettet der ligger gemt i databaser. På overfladen er der pænt og tilgængeligt. På den anden side ligger al slags ufordøjet data, menigsløse tekster, lister og alle mulige ting og hændelser og så en hel del mediefiler. Her ligger der data fra vejrmålinger, informationer om landkort og Dankort, og hemmelige politidata. Man kan sammenligne nettets overflade med vores egen bevidsthed og tankestrøm, mens den gemte side er en slags internetunderbevidsthed. I nettets underbevidsthed findes også "forbudte" tanker, som piratmusik, film og spil og ikke mindst, den mere vilde del af pornoen.

Teknisk kalder man nogle gange den gemte del af nettet for "Deep Web", med en oplagt henvisning til dybhavet. Vi surfer normalt på overfladen, mens der skal mere avanceret udstyr til at komme ned i dybet. De fleste databaser kræver nemlig, at man forstår deres opbygning og har den nødvendige specialsøgemaskine, samt ikke mindst at man har adgangskoden. En meget stor del af den moderne musik ligger i sådanne databaser, som er mere eller mindre tilgængelige via nettet. Nogle af dem er lovlige, for eksempel Apple Computers megasucces iTunes, eller internetboghandlen Amazon, hvor man begge steder kan betale sig til at høre musik.

Faktisk er der også helt stille i mange af de virtuelle musikbutikker. Andre filer er mindre lovlige og tilhører nettets skyggeside af mediekoster og hælervarer. De utallige internetradiostationer befinder sig i en gråzone. En del af dem er helt tilgængelige, og man kan både lytte og gemme musikken med god samvittighed. Andre, som for eksempel DR, kan man kun få lov at lytte på, man må ikke gemme musikken til senere, uden at bryde loven.

For nylig besøgte den amerikanske taleforsker John H.L. Hansen Danmark og brugte lejligheden til at præsentere SpeechFind, som giver brugeren mulighed for at søge efter ord og vendinger i officielle amerikanske taler. Ikke som Google, der søger i internettets tekster, men netop en søgning i den talte lyd. SpeechFind søger i en deep web database: "The National Archive for the Spoken Word".

Talegenkendelse er bare en af de nye lydteknologier, som nu er moden nok til at blive brugt af almindelige mennesker. Et talegenkendelsesprogram oversætter talestrømmen til en tekst, og på tekstform er det lige til at bruge standardsøgemaskiner til at finde interessante vendinger. Med disse nye søgeteknologier er der naturligvis også åbnet for nye skyggesider. Faktisk er den mest omfattende forskningsaktivitet inden for automatisk analyse af tekst, den såkaldte TREC-konference, løbende blevet finansieret af det amerikanske efterretningstvæsen, CIA, som har åbenlyse interesser i at følge med i tale og tekststrømme.

Med denne teknologi åbner der sig dog også betydelige perspektiver for både den private og den professionelle netbruger. Man kan forestille sig, at man logger sine samtaler, ikke bare dem, der føres på telefonen, men også de almindelige samtaler vi fører i løbet af dagen. De almindelige samtaler kunne, for eksempel, gemmes via mikrofonen i headsettet, eller ved at sætte mikrofoner på det nye smart-tøj. Computeren i mobiltelefonen kunne så følge med i lyden, gemme og analysere, eventuelt give feedback til brugeren på baggrund af talens genkendte indhold.

Man taler om "det sociale våbenkapløb": Den der har de bedste it-systemer vil kunne bruge informationen til at overbevise sine kunder og vinde ordren, yde den bedste service, eller simpelthen bare få sin vilje. Det digitale netværkssamfund bygger på arbejdsdeling og på forhandling. Forhandlinger kan afgøres med rå muskelkraft. Men tit er det uklogt at spille for kraftigt med musklerne, man vil jo også gerne kunne se sin forhandlingspartner i øjnene næste morgen. Men forhandlinger kan også vindes med viden, og hvis man er bedre informeret kan man komme ud af en forhandling som vinder, eller endnu bedre: Sørge for at begge parter vinder og kan handle videre til fordel for begge.

Der er en kraftig vækst i den mængde af lyd, som per dag når frem til et almindeligt menneskes øre. Fra tidernes morgen er vores ører designet til naturens lyde, og der var ofte ret stille derude. Vinden i piletræerne og en rislen i åen. En gren, der knækker, er et signal om fare. Og vi farer stadig sammen, hvis vi hører en pludselig uventet lyd, som i reklamerne på nettet. Nogle forskere mener at udviklingen af tale, var baggrunden for at mennesket kunne udbrede sig langt ud over sin "naturlige" udviklingsmæssige niche. Vores evne til at kommunikere er baggrunden for den hypereffektive arbejdsdeling, som siden har erstattet den biologiske evolution med den menneskelige kultur.

Lyd har nogle betydelige fordele som personligt kommunikationsmiddel, frem for eksempelvis, billeder på en computerskærm. Vi skal bruge øjnene, når vi ser på billeder, og det betyder at den visuelle opmærksomhed bliver brugt. Og det ses! Hvis du sms'er, mens du taler med en anden, bliver det bemærket. Men derimod kan man godt forestille sig at du - ubemærket - kunne have lyd i øret. Lyd kan også bruges effektivt til at organisere en gruppe. I Tour de France har samtaleanlæg i flere år været brugt til at sikre, at holdlederen kan ændre på taktikken undervejs i realtime og dermed vinde cykelløb.

Det vakte en del opsigt for nylig, da det kom frem at anføreren for den spanske mesterklub Real Madrid, Raúl Gonzalez Blanco, blev styret af træneren gennem et høreapparat i en pokalfodboldkamp. Lyd er en meget vigtig del af moderne fodbold. Kommunikerende spillere, som danske Gravesen, gør en forskel. De kan sørge for at taktikken holdes og indirekte være trænerens højttaler.

På DTU's institut for informatik og matematisk modellering, IMM, forsker vi i intelligente lydssystemer. Computere bliver stadig bedre til at forstå lyd. Vi kan i dag lære en kunstig intelligent computer, at lytte til musik og for eksempel hjælpe musikbrugeren med at genrebestemme musik. Det betyder at man kan hente musik fra de mange radiokanaler og sortere den efter ens individuelle smag og behov. Sammen med kolleger fra Aalborg Universitet har vi etableret et lydforskningsprojekt, der er finansieret af Forskningsrådet for Teknik og produktion. Projektet skal forberede os for fremtidens lydforbrug. Det overordnede mål er bedre lydsoegemaskiner, der kan hjælpe os med administrere det kolossale medieudbud.

Vi forsøger for eksempel at træne computere til at adskille tale og musik i nyhedsudsendelser og til at genkende talere og kunstnere. Når computeren på den måde har forstået udsendelsens indhold, kan den nedskrive den talte information og genrebeskrive udsendelsens musik. Dermed får vi mulighed for at søge efter

oplysninger i radioudsendelser, som vi i dag søger i internettets tekst.

Der er især gjort store fremskridt i de sidste par år inden for adskillelse af blandede signaler. Generelt kalder man det "cocktailparty-problemet": At der er mange der taler i munden på hinanden. Vores hjerne er rigtig god til denne adskillelse, mens det tidligere har været svært at få computer til at lære det. En koncentreret forskningsindsats gennem de senere år har nu ledt til brugbare løsninger for lyd, og vi forsøger i øjeblikket at generalisere løsningerne til alle typer af blandede medier. I en verden, hvor vi stadig oftere befinder os i lyd- og billedintense miljøer, vil vore medfødte evner til at fokusere på det vigtige og sortere resten fra, blive alvorligt udfordret. Men vi håber at de nye metoder og computerprogrammer vil kunne hjælpe os til at navigere sikkert i de nye medier.

Vi lever i dag i en slags stilhed før fremtidens kakofoniske mediekaos og der er et stort behov for systemer, der kan hjælpe os med at administrere de mange muligheder, og måske endda ind imellem skabe lidt velfortjent hvile for hårdt prøvede ører og øjne.