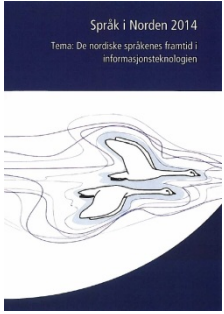


# Sprog i Norden

|            |   |   |
|------------|---|---|
| Titel:     | Tale er Guld<br>Om talegenkendelse i de danske kommuners tjeneste   |  |
| Forfatter: | Peter Juel Henriksen  |   |
| Kilde:     | Sprog i Norden, 2014, s. 24-32  |   |
| URL:       | <a href="http://ojs.statsbiblioteket.dk/index.php/sin/issue/archive">http://ojs.statsbiblioteket.dk/index.php/sin/issue/archive</a> |   |

© Forfatterne og Netværket for sprognavnene i Norden

## Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

## Søgbarhed

Artiklerne i de ældre numre af Sprog i Norden (1970-2004) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

# Tale er Guld

## Om talegenkendelse i de danske kommuners tjeneste

*Peter Juel Henriksen*

*Mange danske kommuner er parate til at indfase automatisk talegenkendelse, men er samtidig nervøse efter en lang række dårlige businesscases i den nærmere fortid. Der klages over høje licenspriser og lavt serviceniveau, den typiske virkning af et de facto monopol på leverandørsiden. Denne artikel beretter om en ny offentlig handlingsstrategi der har taget form i 2013. Nok er Danmark et lille sprogområde, men vores offentlige sektor er velorganiseret og købedygtig. Det seneste år har de danske kommuner organiseret sig omkring den taleteknologiske udfordring, man har i fællesskab formuleret en række spørgsmål og nye krav til leverandørerne og bevidst arbejdet frem mod et udbudsmateriale der åbner for fair konkurrence. I dette arbejde har også offentlige forskere, heriblandt denne artikels forfatter, deltaget som økonomisk neutrale rådgivere. Initiativet skal bane vej for den første rentable kontrakt på området - som vi håber at se i 2014 - en begivenhed som præcis kunne løsne den nuværende deadlock og åbne et milliardstort marked for taleteknologi i det offentlige tjeneste.*

De danske hospitaler er erfarne brugere af talegenkendelse. De danske kommuner ville meget gerne være det også. Set fra en kommunalbestyrelses synspunkt har talegenkendelse en række indlysende fordele; den kan gøre sagsbehandleren mere produktiv, fjerne en forhadt skrivebyrde for de praktisk orienterede funktionærer, støtte de sprogligt udfordrede medarbejdere: ordblinde, nydanske, svagsynede, lavtuddannede. Med talegenkendelse på sin smartphone eller tablet kan hjemmehjælperen afslutte sin visitationsrapport undervejs til næste klient, og trafiktælleren føre sine observationer direkte ind i regnearket. Kort sagt, talegenkendelse er *lean*.

## Baggrund

En lille snes kommuner - ud af Danmarks otteoghalvfems (98) - har sat penge i taleteknologiske produkter gennem de sidste fire-fem år. De har generelt været seriøst indstillet på at forberede deres personale på nye arbejdsrutiner hvor man dikterer frem for at taste; ingen har troet at fordelene skulle komme uden samvittighedsfuldt forarbejde. Alligevel har endnu ingen dansk kommune tjent sin investering hjem. I dag anvender kun ganske få kommunalt ansatte talegenkendelse på daglig basis, hvad der står i skarp kontrast til de titusindvis af læger i Skandinavien der benytter teknologien. Problemet har især ligget i tilpasningsfasen. De fleste kommuner giver udtryk for at deres leverandør lovede en høj service med uddannelse af brugere og tilpasning af talegenkendelsen *in situ*, men at virkeligheden blev en anden efter at kontrakten var indgået og leverandøren efter kort tid overlod driften til brugerne selv. En anden kilde til irritation er de meget lange perioder man må vente på gentrænede systemer (fx sprogmodeller og akustiske modeller optimeret til de aktuelle arbejdsforhold). Producenterne hævder, på deres side, at de ikke kan tjene penge på at levere talegenkendelse til den danske kommune i dag, og at man kun er gået ind i markedet af strategiske grunde.

På softwaresiden har de danske kommuner reelt kun haft én teknologi at vælge imellem, nemlig Nuance Communications. Om denne teknologi er formidlet og tilpasset af skandinaviske firmaer som KMD, IBM Denmark, Max Manus eller PDC/Dictus forandrer ikke det faktum at licensbetingelserne er dikteret fra USA. På grund af Nuance's de facto-monopol har kommunerne haft svært ved at lægge pres på deres direkte leverandører, da disses margin for konkurrence er uhyre smal og, som nævnt, styret af international kapital. At dømme efter Nuance's nyere PR satses koncernen i de kommende år på intelligente dialogløsninger rettet mod (primært) amerikanske firmaer og velhavende forbrugere, frem for at tilgødese de europæiske markeder og de mindre sprogområder. Denne strategi er for nyligt blevet accentueret da Nuance antog professor Ron Kaplan som øverste chef for sin forskningsafdeling, en verdenskendt forsker i formel grammatik og kunstig intelligens - og ikke i små sprog.

## Nye initiativer - kommuner og universitet i samarbejde

Hvordan bevæger de danske kommuner sig i dette minefelt? De kommunale budgetter er under pres i disse kriseår, og mange borgmestre skal vælge mellem fyringer og effektiviseringer. De spiller gerne bolden vi-

dere til deres IT-ansvarlige, med besked om at udrette et teknologisk mirakel som på én gang kan støtte de svageste ansatte og øge kommunens produktion. Principielt er dette jo ikke umuligt, men det kræver nye tekniske kompetencer hos de IT-ansvarlige, frigørelse fra gamle og tyngende industrikontrakter, vidensudveksling og genbrug af træningsmaterialer over kommunegrænser, *esprit mousquetaire*. Derfor er en lang række af de danske kommuner gået sammen i et frivilligt fællesskab, kaldet OS2, som mødes rundt i landet og udveksler erfaringer om indfasning af ny teknologi. OS2 har allerede haft succeser med yderst positive businesscases, og OS2-kredsen vokser derfor hurtigt og tæller nu omkring halvdel af de danske kommuner. Da taleteknologien er et særlig varmt emne, har de meste handlingsparate kommuner i OS2 dannet undergruppen *OS2Talk*, der gennem 2013 har taget initiativer til at bringe de centrale aktører sammen: De danske leverandører af taleteknologi (nuværende og potentielle), kommunerne selv, KOMBIT<sup>1</sup>, Digitaliseringsstyrelsen – desuden kernebrugere i hospitalsvæsenet, Folketinget og Danmarks Radio.

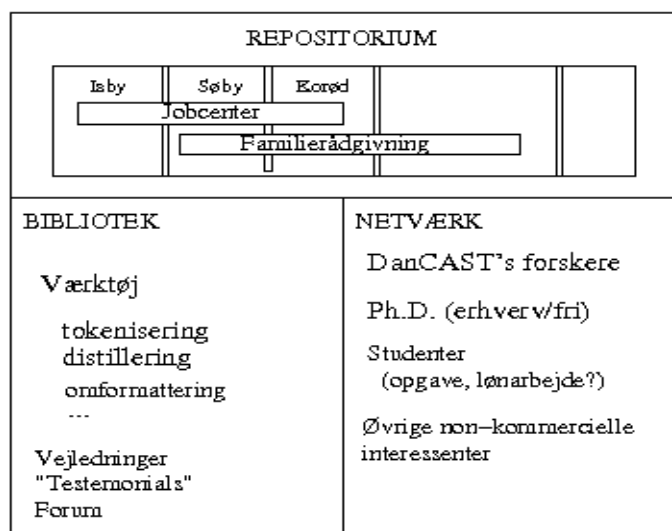
De fleste af stormøderne har været afholdt på Copenhagen Business School (CBS) på Frederiksberg, organiseret af forskningscenteret DanCAST. Dette forskningscenter har i de seneste år samlet sine aktiviteter om to temaer, dels de særlige udfordringer som det danske (og de øvrige nordiske) sprog stiller til taleteknologien, og dels det anvendte perspektiv: hvordan dagens teknologiske løsninger tilpasses et mangeartet brugsmiljø, hvor stemmetyper, regiolekter, ordforråd, arbejdsvaner, støj og rumklang giver praktiske udfordringer som grundforskeren ofte abstraherer væk i sit laboratorium. For at gøre vejen så kort som muligt fra vores forskning til samfundets arbejdspladser har DanCAST åbnet et særligt forum som betjener kommuner med en dagsorden inden for taleteknologi. Dette forum kalder vi *The DanCAST hub for ASR Localization Support*.

Teknisk set er DanCASTs hub en stor server stillet til rådighed af CBS's IT-Afdeling, med garanti for vedligehold og opetid i alt fald fem år ud i fremtiden. Til serveren hører et lager med garanteret 500GB diskplads til hver af landets kommuner (på længere sigt er det IT-Afdelingens plan at lægge repositoret "i skyen").

---

1 KOMBIT er et selvejende (og derfor økonomisk uafhængigt) aktieselskab som rådgiver de danske kommuner om alle aspekter af IT-indkøb og anvendelse, lige fra udarbejdelsen af fair og åbne udbudsmaterialer, over efteruddannelse af personale, til juraen omkring misligholdelse af kontrakter.

### The DanCAST hub for ASR localization support



DanCASTs HUB ("The DanCAST hub for ASR localization support")

Enhver dansk kommune med en ASR-dagsorden (automatic speech recognition) kan tegne en samarbejdsaftale med DanCAST. Aftalen forpligter CBS til at åbne og vedligeholde et repositorium hvortil kommunen har lov at uploade et stort materiale af tekster. Disse skal være udvalgt sådan at de støtter den træning som udviklingen af en ASR-løsning bygger på. Når de danske kommuner samler deres træningsdata på ét sted, giver det mulighed for at organisere dataene i adskillige fagområder, fx Jobcenter, Ældreområdet, Børn & Ungdom, uden at datamængden bliver for lille i de enkelte sektorer. Som datamængden vokser, bliver det lettere og lettere for den sidst tilkomne kommune at opnå et effektivt træningsgrundlag ved blot at tilføje sit eget mindre bidrag af dokumenter. Det er ingen

hemmelighed at den gode genkendelsesgrad i et ASR-system afhænger af et træningsmateriale der præcist dækker det ordforråd som applikationen skal genkende.

DanCASTs hub vedligeholder også et bibliotek af software, testemonials og andre dokumenter. Herfra kan kommunen downloade scripts og manualer og selv påbegynde arbejdet med at præparere sine tekster. Før et tekstkorpus kan anvendes til ASR-træning, skal det nemlig gennemgå en lang række processer til oprensning og normalisering.

- Formatering (dokumenter omlægges fra fx SQL, regneark, eller docx til klar tekst)
- Frasortering (redundant tekst, gentagelser, irrelevant markup, sidefod og sidehoved, ...)
- Tokenisering (ord frigøres fra omgivende tekstelementer såsom kommaer, parenteser, ...)
- Normalisering (forkortelser, numre, beløb, tider, etc. ændres til bogstavede former)
- Anonymisering (følsomme informationer skjules: navne, cpr-numre, steder, etc.)
- Annotation (informationer af særlig betydning mærkes op)

Herunder er vist et stykke tekst fra en kommunal visitationsrapport. Fra den øverste form til den nederste form har teksten gennemgået en tokenisering, normalisering og anonymisering (forskellene er markeret med **fed** font; personnavn og dato er opdigtet). Teksten er derved blevet forberedt til træning af et ASR-systems sprogmodel.

---

\* Prognose og behandlingsmuligheder (**1-3-12**):

Der er klar **overenstemmelse** mellem de objektive fund og borgers fortælling..

- Foreligger der lægeskøn eller **udtalelse** fra egen læge **ang. pronose, arb evne mm?**

Ja. Egen læge ser ikke noget arbejdsmarkedsperspektiv for **Marianne**.

---

prognose og behandlingsmuligheder  
første i tredje tolv

der er klar overensstemmelse mellem de objektive fund og borgers  
fortælling

foreligger der lægeskøn eller udtalelse fra egen læge angående pro-  
gnose arbejdsevne med mere ja egen læge ser ikke noget arbejdsmar-  
kedsperspektiv for \_FirstName\_

---

Til hvert af de nævnte trin i tekstforberedelsen stiller DanCASTs hub værktøj til rådighed (ikke-proprietært), sådan at kommunen får et reelt valg om den ønsker at overtage disse opgaver selv eller overlade dem til en totalentreprenør i en udbudssituation. Dette giver en ny valgfrihed, idet alle kontrakter mellem kommuner og leverandører til dato har placeret den samlede udviklingsopgave hos leverandøren. For nogle kommuner har det været en grim overraskelse at opdage at de derved har mistet rettighederne til egne data, som efter den nævnte forbehandling er overgået til ASR-leverandøren.

Endelig, som sit tredje indsatsområde, vedligeholder DanCASTs hub et kartotek af personer og organisationer med en særlig interesse i området (som for eksempel den centrale IT-ansvarlige i hver af kommunerne). Desuden giver vi små og mellemstore danske firmaer lov til at gøre opmærksom på sig selv, med en kort præsentation og kontaktinformation. På denne måde kan kommuner som måtte ønske at tage de første skridt i et ASR-projekt forud for en udbudsrunde, let finde hjælpende hænder med passende kompetencer, enten det er et lille nystartet firma eller en lingvistisk studerende på mindsteløn.

CBS tager ikke penge for sine ydelser, og DanCASTs tilbud til kommunerne er derfor gratis. Til gengæld giver kommunerne, gennem samarbejdsaftalen, CBS lov til at forske (not-for-profit) i de uploadede data, naturligvis sådan at CBS altid skal respektere kommunernes almene copyrights og den enkelte kommunes særlige anonymitetskrav. Desuden skal CBS slette alle kopier af lånte data, hvis en kommune vælger at afslutte sin aftale med DanCAST. Selv om DanCASTs hub kun har været åben for samarbejder i få uger, har allerede fem kommuner afleveret tekstdata til vores repository. Vi forventer lige så mange nye aftaler inden jul.

## Høringen i november 2013

Om få uger indtræffer den store høring mellem kommuner og leverandører. Dybest set er høringens formål at sikre ASR-projekter der giver overskud for både køber og sælger, for hvis ikke vi snart opnår positive businesscases, risikerer vi at ASR ender helt uden for kommunernes interessefære.

Et vellykket udviklingsforløb kræver en detaljeret problemafklarung *før* kommunen går i konkret udbud. En stor del af afklaringsprocessen er fælles for alle kommuner, og tanken med høringen er at få tidlige svar fra industrien på så mange udfordringer som muligt. På den baggrund vil kommunerne stå bedre rustet til at udarbejde et generisk udbudsmateriale, som den enkelte kommune derefter kan tage som udgangspunkt.

For at få så megen nytte af høringen som muligt har en lønnet projektgruppe i månedsvi arbejdet med et oplæg, som nu er rundsendt til de potentielle ASR-leverandører. Oplægget konkretiserer de temaer som kommunerne i OS2 har kredset om i årevis. Hvilke krav kan og skal en kommune stille i sit udbudsmateriale? Hvilke dele af en ASR-løsning kan kommunen selv have ansvaret for at udvikle? Hvilke moduler i en samlet løsning kan kommunen bevare ejerskab til og genbruge i en senere udbudssituation (når der måske har vist sig en ny og stærkere taleteknologi i markedet)? Blandt de næsten hundrede punkter venter vi særligt spændt på industriens reaktioner på disse:

- Tilbudsgiver skal redegøre for hvor hyppigt den akustiske model opdateres.
- Tilbudsgiver skal redegøre for om det er muligt for ordregiver at få den akustiske model opdateret efter ønske.
- Det skal være muligt at anvende mere end en sprogmodel i den tilbudte løsning, for eksempel en sprogmodel til beskæftigelsesområdet og en sprogmodel til miljøområdet.
- Det skal være muligt at få gentrænet eksisterende sprogmodeller efter ønske.
- Tilbudsgiver skal redegøre for hvor lang tid det tager at gentræne en eksisterende sprogmodel.



- Tilbudsgiver skal redegøre for hvilke leverancer ordregiver i givet fald selv kan stå for, ved udarbejdelse af nye sprogmodeller.
- Tilbudsgiver skal redegøre for hvilke trin i forædlingsprocessen [tekstforberedelsen omtalt herover] kommunen selv kan udføre og hvilke tilbudsgiver skal udføre.
- Tilbudsgiver skal redegøre for om der er mulighed for at genbruge eventuelt udviklede databaser (for eksempel den fonetiske ordbog) ved senere inddragelse af nye taleteknologier som Text-to-Speech.

Det samlede høringsmateriale kan beses ved henvendelse til forfatteren.

## Afsluttende bemærkninger

Den nyligt udkomne hvidbog, Europe's Languages in the Digital Age, konkluderer at de fleste europæiske sprog er dårligt forsynet med sprogteknologi.<sup>2</sup> Blandt de europæiske lande udmærker Norden sig ikke særligt. Vores sprog er for små til at hjemmemarkederne af sig selv genererer firmaer og produkter til sine egne sprog. Det gælder især de tungere teknologier som talesyntese (text-to-speech), automatiske oversættelse og talegenkendelse, hvor man (med få undtagelser) kun kan håbe på produkter fra internationale leverandører der i bund og grund er tillempede løsninger konciperet til andre sprog. Kvaliteten er næsten aldrig på niveau med den tilsvarende engelske, hvad der giver vores yngre generation endnu en grund til at svippe over i engelsk når de skal betjene deres hardware, browsere, GPS'er, biler, og så videre. Denne udvikling er blevet kaldt *den digitale sprogdød*.

Vi tror ikke at der er grund til dødsangst. Nok er vores sprog små, men af langt større vigtighed end det nøgne antal af talere er markedets grad af organisation og stabilitet. I Norden vil den offentlige sektor til alle tider være en af samfundets stærkeste og mest stabile købere. Når først de vellykkede projekter begynder at dominere, bliver der plads til et helt økosystem af leverandører. Et naivt håb? Næppe. Alene rygtet om mere fair udbudsrunder har allerede fået to nye danske firmaer til at tjekke ind blandt tilbudsgiverne. Måske synger monopolet på sit sidste vers.

---

<sup>2</sup> <http://www.meta-net.eu/whitepapers/key-results-and-cross-language-comparison>

## **Organisationer og andre der er nævnt i artiklen:**

|                          |   |
|--------------------------|---|
| OS2                      | <a href="http://www.os2web.dk">http://www.os2web.dk</a>                                   |
| OS2Talk                  | <a href="http://www.os2web.dk/document/os2talk">http://www.os2web.dk/document/os2talk</a> |
| KOMBIT                   | <a href="http://www.kombit.dk">http://www.kombit.dk</a>                                   |
| Digitaliseringsstyrelsen | <a href="http://www.digst.dk">http://www.digst.dk</a>                                     |
| CBS/language-technology  | <a href="http://bridge.cbs.dk">http://bridge.cbs.dk</a>                                   |
| DanCAST                  | <a href="http://www.DanCAST.dk">http://www.DanCAST.dk</a>                                 |

**Peter Juel Henriksen** er associate professor og principal investigator ved Danish Center for Applied Speech Technology (DanCAST), Copenhagen Business School, [pjh.ibr@cbs.dk](mailto:pjh.ibr@cbs.dk)

## **Summary**

Many Danish municipalities are quite ready to implement automatic speech recognition, however at the same time reluctant after a line of negative business cases in the recent past. Complaints have been heard of over-priced licenses and low service level, the typical effect of a de facto monopoly situation in the ICT market. This article presents a new public effectuation strategy installed in 2013. Denmark may well be a small country, but our public sector is highly organized and (for this reason) economically quite capable. During the last 12 months the municipalities have organized themselves, formulating together a list of requirements to be met by the suppliers as a part of the new tender documents. This work has been supervised by university based researchers (including this author) acting as commercially unbiased advisors. The initiative is intended as a significant move towards the first mutually profitable ASR contract in this area, an event we hope will occur in 2014 and which may be exactly what is needed to dissolve the current deadlock and open up a billion kroner market for ASR in the Danish public sector.