

Et effektivt teknologivalg til sundhedssektoren

Talegenkendelse er allerede indført på flere danske sygehuse over hele landet. Brugere af talegenkendelse er glade for teknologien og peger på flere fordele ved talegenkendelse, hvor både behandlingshastigheden og behandlingskvaliteten øges, når talegenkendelse indføres.

Max Manus A/S benytter taleteknologien **SpeechMagic™** fra Nuance udviklet af **Philips**. SpeechMagic™ genkender det talte ord og konverterer det til et tekstdokument.

SpeechMagic™ kan anvendes som en selvstændig teknologi, men det er specielt, når det anvendes integreret i andre systemer, fx EPJ, Røntgen eller patologisystemer med **SpeechEditor™**, at de største gevinster opnås!

Opbygning af SpeechMagic™

SpeechMagic™ systemet er bl.a. opbygget af en **recognition server** og en **adaptation server**. **Recognition serveren** danner tekstdokumentet og står for den akustiske indlæring, hvorimod **adaptation serveren** "lærer" af de rettelser vi laver og står for indlæring af sprogmodellen.

Online- eller batch-diktering.

Lægerne kan vælge enten at diktere **online**, det vil sige, hvor ens tale bliver konverteret til tekst med det samme, (oversættelsen fra tale til tekst foregår på den lokale klient), eller at **batch**-diktere.

Batch-diktering minder om traditionel diktering. Den dannede lydfil sendes til genkendelse via recognition serveren. Den genkendte tekst vil herefter kunne kaldes frem (evt. af en sekretær) og rettes.

Efter **online**-diktering eller rettelse af **batch**-diktering sendes teksten med eventuelle rettelser til en adaption server, som husker de rettelser, man har lavet.

Systemet vil så langsomt lære den enkelte bruger at kende eller rettere vedkommendes stemme og udtalemåde (fonetiske profil) og opbygning af sætninger.

Oplæring med talegenkendelse

Talegenkendelse betyder en stor omvæltning for både lægerne og sekretærer.

Lægerne skal diktere på en ny måde. Lægen skal diktere tegnsætning, nye linjer og nye afsnit. Man kan ikke komme med små bemærkninger til sekretæren om at tilføje eller at slette noget. Lægerne skal være meget opmærksomme på, hvordan de udtaler ordene og undgå lange pauser. Hvis man skal bruge en tænkepause, skal man standse dikteringen. Man skal **udtale ordene tydeligt og** undgå at mumle eller at tale for hurtigt. Det er ligeledes vigtigt, at man ikke sletter/retter unødigt i den genkendte tekst.

Der er ikke kun tale om, at systemet skal lære de enkelte brugere at kende. Lægerne skal også lære systemet at kende. Hvis talegenkendelse skal blive en succes, kræves der ikke kun, at teknologien fungerer, men også at læger og sekretærer er motiverede.



Teknikken bag SpeechMagic™

SpeechMagic™ leveres med en kontekst (Context=ordbog), der indeholder tusindvis af ord og begreber brugt inden for det kliniske område. Det betyder, at SpeechMagic™ allerede fra starten kan genkende de fleste fagudtryk, der anvendes på danske kliniske institutioner (sygehuse, læge-praksis mv.).

Denne kontekst består af et **leksikon**, som indeholder ca. 65.000 ord, en **sprogmodel** og noget grammatik, der bl.a. angiver, hvordan man skal sige f.eks. datoer (jf. "Dictation Guidelines").

Leksikonet indeholder selve ordbogen, der kan være fælles for alle læger, - f.eks i en afdeling. I ordbogen er hvert ord repræsenteret i almindelig skrevet form samt i form af lydskrift, hvor hvert lydskriftsymbol svarer til en akustisk model.

Der er endvidere en række 'flags' i ordbogen, der f.eks. angiver, at er der genkendt et "punktum", så skal næste ord skrives med stort. I leksikonet er der også informationer om, hvornår ordet sidst har været brugt, og hvor ofte det bruges. Efterhånden som brugerne af systemet tilføjer ord til ordbogen, så sker der på et tidspunkt en omorganisering af leksikonet således, at de ord man aldrig bruger inaktiveres (baggrundsleksikon).

Første gang man skal benytte systemet, skal det introduceres til brugerens stemme. Man laver således et akustisk "fingeraftryk". Dette sker første gang ved hjælp af programmet "Initial training".

Den akustiske profil, som er brugerspecifik, registrerer, hvordan taleren udtaler sprogets dellyde.

Konteksten indeholder ligeledes **en sprogmodel**. Denne model er specifik pr. taler ligesom det akustiske fingeraftryk. I denne model lagres statistisk information om, hvordan netop denne taler typisk opbygger sine sætninger. Den indeholder simpelt hen sandsynligheder for, at et givent ord følger et andet.

Fra tale til tekst

Når talen omsættes til tekst, matcher man lydbilledet op imod først de akustiske dellyde - man finder den streng, der matcher lydbilledet bedst, og ser herefter i leksikonet om man har et ord med lydskrift, der svarer til denne "dellydsstreng".

Derefter anvendes sprogmodellen til at finde den mest sandsynlige streng af ord, og når disse sandsynligheder tilsammen er optimeret, har systemet fundet "den mest sandsynlige tekststreng", som talen skal oversættes til. Herved er talen omsat til tekst.

Mikrofonen skal ligeledes indstilles til de lydforhold, som man arbejder i, og hvor højt man taler. Dette gøres ved hjælp af en Audio Wizard.

I den efterfølgende anvendelse af SpeechMagic™ sker en løbende successiv læring af både den akustiske profil og sprogmodellen, hver gang der dikteres.

Når en genkendt tekst er rettet, sammenlignes genkendt og rettet tekst fortløbende og eventuelt nye ord, som systemet ikke kender, tilskrives en liste. Listen af ukendte ord (**Unknown Words**) håndteres af en superbruger via SpeechMagic Context Adaptation modulet. Dermed tilføjes der løbende nye ord til leksikonet, som således udvides og tunes mod brugen på netop det aktuelle afsnit/afdeling.

"Ordbogen" udbygges således af de personlige rettelser, der bliver udført i oversat tekst, hvorved der over en periode vil blive opbygget et betydeligt ordforråd. Dermed være sagt, at jo mere der dikteres, des bedre bliver genkendelsen. Der kan ikke garanteres en given genkendelses-procent. Mange faktorer kan påvirke genkendelsen. Bl.a. dikteremåde, diktat- og rettedisciplin, omfanget af rettelser, yderligere tilføjelser/slettelser mv. Erfaringen siger, at den bedste genkendelse opnås ved, at lægen retter selv og dermed får direkte feedback på dikterestilen.

Den enkelte persons genkendelsesgrad vil således være *meget* afhængig af, hvordan der dikteres, og hvor disciplineret, der dikteres og eventuelt rettes.

Kontakt os for at få en demonstration eller yderligere oplysninger omkring de løsningsmuligheder, vi kan tilbyde på tlf. 70 10 55 10