

*Tomi Kauppinen, Mika Koivikko,
Seppo Koskinen, Juhani Ahovuo*

Puheentunnistus nopeuttaa lausuntojen valmistumista

Tärkein tieto

- › Ensimmäinen suomenkielinen erikoissairaanhoidon puheentunnistusjärjestelmä luotiin radiologian alalle.
- › Kutakin erikoisalaa ja kielimallia varten tarvitaan oma erikoissanasto, jonka avulla puheentunnistuksen tunnistustarkkuutta saadaan oleellisesti parannettua.
- › Puheentunnistuksen käyttäminen nopeuttaa lausunto-prosessin läpimenoaikaa. Lausunto on kliinikon nähtävissä välittömästi hyväksymisen jälkeen.
- › Puheentunnistuksen käyttöönotto on varsin helppoa, sillä uusi käyttäjä tarvitsee keskimäärin vain 10–15 minuutin perehdytyksen.
- › Puheentunnistuksen käyttäminen parantaa lausuntojen informatiivisuutta, rakennetta ja selkeyttä.

Puhuessa tuotettu ääni on ilmanpaineen nopeaa värähtelyä, jonka tunnistaminen on ihmiskorvalle varsin helppoa, mutta koneelle vaikeata. Puheentunnistuksessa puhe muutetaan digitaaliseksi signaaliksi, josta erilaisin menetelmin käsittelemällä muodostetaan tulkintoja kielellisten ilmausten todennäköisyyksistä.

Puheentunnistuksella tarkoitetaan sovellusta, joka pystyy tunnistamaan tuotetun puheen puhujasta riippumattomasti ja tuottamaan sen suomenkieliseksi sanoiksi ja tekstiksi. Sovellus on tietokonepohjainen, ja se muuttaa sanellun puheen tekstiksi automaattisesti. Tämä merkitsee sitä, että erillinen tekstinkäsittelijän suorittama konekirjoitusvaihe jää pois ja kokonaisprosessi nopeutuu ja tehostuu.

Yksinkertaisin puheentunnistuksen muoto perustuu komentokieleen; sitä sovelletaan esim. matkapuhelimien puheentunnistuskomennoissa. Toinen puheentunnistuksen

muoto perustuu kommunikaatioon, ja sitä hyödynnetään mm. asiakaskeskuksien rutiinitehtävissä, erilaisissa tietokantakyselyissä sekä tietokoneen puheohjauksessa. Kolmas ja vaativin puheentunnistuksen muoto on täysin vapaa puheentunnistus, jota voidaan hyödyntää mm. lausuntotekstien ja sanelujen käsittelyssä.

Suomen kieli on puheentunnistukselle haaste, koska siinä on lukuisia sijamuotoja ja verbeillä erilaisia päätteitä. Laajasakaan sanavarastossa ei voi olla kaikkia sanoja kaikkine taipumusmuotoineen. Puheentunnistusjärjestelmän tehtävänä on liittää yhteen sanavartalot ja -päätteet. Puheentunnistuksella voidaan sanella tekstejä, mutta vapaan sanelun laaja sanasto edellyttää teknisesti tehokasta laitekapasiteettia ja kehittyneitä tunnistusmenetelmiä.

Puheentunnistuksen käyttökohteita

Puheentunnistusjärjestelmiä alettiin kehittää jo 1980-luvulla, mutta vasta 1990-luvun lopulla sovelluksista tuli niin kehittyneitä, että ne alkoivat yleistyä (1,2). Alussa puheentunnistusta käytettiin yleiseurooppalaisissa kielissä, kuten englannissa, saksassa ja ranskassa, mutta nykyisin puheentunnistus toimii useilla muillakin kielillä.

Puheentunnistusohjelmia on kehitetty sekä koti- että toimistokäyttöön ja lisäksi vaativaan kliiniseen käyttöön. Menetelmän etuja on yritetty hyödyntää laaja-alaisesti, esimerkiksi peruskäyttöisissä tietokoneissa, ja mm. Microsoftilla on XP-käyttöjärjestelmään integroituna oma puheentunnistussovellus, jolla voi sanella dokumentteja Microsoft Office -ohjelmiin. Myös muilla ohjelmistovalmistajilla on omia sovelluksia peruspuheentunnistukseen, esim. Philips FreeSpeech Viva, Lingsoft sekä IBM ViaVoice Pro ja WebSphere Voice.

Lääketieteellisiä lausuntoja ei tällaisilla ohjelmilla kuitenkaan pystytä riittävän luotettavasti tuottamaan, koska se edellyttää jatkuvaa ja vapaata puheentunnistusta. Ensimmäiset suomenkieliset lääketieteelliseen käyttöön suunnatut pu-

heentunnistussovellukset tulivat vasta muutamia vuosia sitten. Toistaiseksi suomeksi on laadittu perusterveydenhuollon sanasto sekä erikoissairaanhoidon aloista radiologian sanasto (3). Puheentunnistus voidaan integroida useimpiin radiologian tietojärjestelmiin (radiology information system, RIS). Lääketieteelliset lausunnot vaativat omat erikoisohjelmansa, joita on mm. Philipsilla ja IBM:lla (2,4). Philipsin kehittämä SpeechMagic on eniten Euroopassa käytetty puheentunnistusratkaisu (4).

Suomen kieli ja erikoissanasto

Puheentunnistus voidaan jakaa osavaiheisiin, jossa ensin puhesignaali muutetaan digitaaliseen muotoon ja saatua signaalia käsitellään. Signaalista pyritään erottamaan jaksot, joissa puhuja puhuu tai on hiljaa. Tässä vaiheessa myös taustakohinaa pystytään suodattamaan pois. Mallinnuksen jälkeen jaksoille pyritään löytämään tilastollinen muoto ja todennäköisimmät mahdollisuudet. Tyypillisesti tähän hyödynnetään ns. Markovin piilomallia (Hidden Markov Model), joka sisältää mallinnuksen itseoppivuutta auttavia tunnistusalgoritmeja (5).

Suomen kieli on puheentunnistusjärjestelmälle haastava, koska sanasto on laaja ja kasvaa nopeasti. Lisäksi yhdestä sanavartalosta on mahdollista muodostaa useita eri sanamuotoja. Kutakin erikoisanaston sanastoa ja kielimallia varten tarvitaan oma erikoissanasto (context). Erikoissanaston avulla pystytään ottamaan huomioon sanojen esiintymistiheys, erilaiset lausemallit sekä uudet alakohtaiset ilmaisut. Erikoissanastoa käyttämällä tunnistustarkkuutta saadaan oleellisesti parannettua, ja siksi sanastot ovat aina kieli- ja alakohtaisia.

Puheentunnistus on oppiva järjestelmä. Tuntemattomat ja tunnistamattomat sanat voidaan lisätä erikoissanastoon, jolloin seuraavan kerran saneltaessa järjestelmä tunnistaa sanan. Järjestelmä siis paranee ja erikoissanasto laajenee kerta kerralta. Myös usein toistuvat tarpeettomat tauot, äänneet ("hmm", "öö") tai kokonaiset lauseet ("Sanelu päättyy tähän, kiitos"), jotka kenties normaalisti sanellaan mutta joita ei lopulliseen lausuntotekstiin kirjoiteta, voidaan erikoissanaston avulla jättää tekstistä automaattisesti pois. Tällöin niitä ei tarvitse poistaa tai korjata erillisen korjausprosessin aikana.

Radiologian toimintaprosessi

Puheentunnistusta käytettäessä radiologi näkee saneleman tekstin muutaman millisekunnin viiveellä ja voi korjata saneluaan käyttäen joko apunaan näppäimistöä tai sanelemalla uudelleen virheellisen tekstin päälle. Radiologi vahvistaa tai allekirjoittaa lausunnon, ja sen jälkeen lausunto on välittömästi klinikon nähtävillä (kuvio 1).

Radiologi voi välittää lausunnon myös tekstinkäsittelijän tarkastettavaksi, jolloin tekstinkäsittelijä kuulee radiologin sanelun ja näkee puheentunnistuksen tunnistaman tekstin. Tekstinkäsittelijän tarkastaessa lausunnon ajansäästö on kuitenkin pienempi, koska lausuntoprosessiin tulee ylimääräinen työvaihe mukaan, vaikka tekstinkäsittelijän ei tarvitsekaan kirjoittaa tekstiä. Tekstinkäsittelijöiden ajankäyttö voi olla hyvinkin yksilöllistä ja vaihtelee myös kokemuksen mukaan.

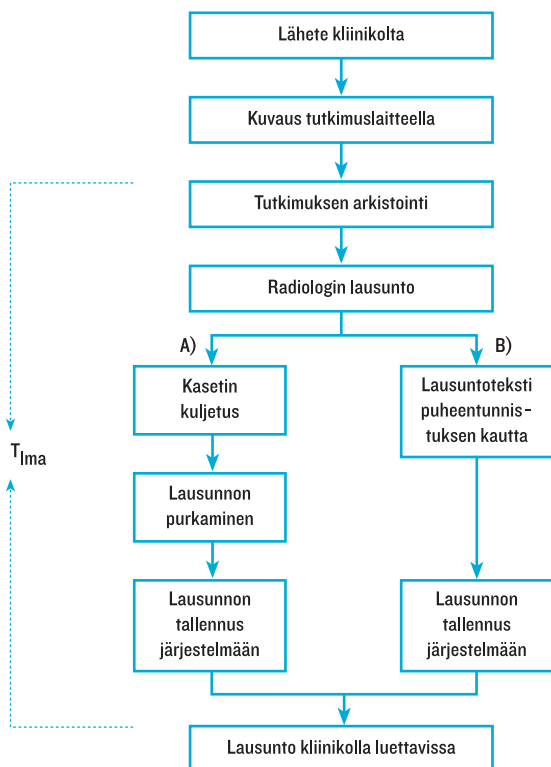
Tekninen toteutus

Puheentunnistusjärjestelmä edellyttää sähköistä potilastietojärjestelmää, johon puheentunnistus on integroitu. Sosiaali- ja terveysministeriön kansallinen terveystietojärjestelmä on tuonut mukanaan vaatimuksen sähköisten potilastietojärjestelmien käyttämisestä (6). Radiologiassa tällaisia valmiuksia ja järjestelmiä alkaa olla jo lähes kaikilla järjestelmätoimittajilla, mutta muualle erikoissairaanhoidon järjestelmät tekevät vasta tuloaan. Puheentunnistus on käytettävissä yhdellä perusterveydenhuollon järjestelmätoimittajalla, mutta käyttö on laajenemassa myös muiden valmistajien tuotteisiin.

Vaikka puheentunnistus itsessään ei edellytä suuria vaatimuksia käytettävältä tietokoneelta, niin radiologisessa työkentelyssä kuvankatselutyöaseman tulee kuitenkin olla varsin moderni ja tehokas. Kokemusten mukaan puheentunnistus toimii joustavasti työasemilla, joissa on keskusmuistia vähintään 1 tai 2 GB. Puheentunnistuksen käytettävyyttä lisää myös käytössä olevien laitteiden kokonaislukumäärä, ja

Kuvio 1.

Radiologinen lausuntoprosessi. A kuvaa perinteistä lausuntoprosessia, jossa sanelu tehdään kasetille, josta se puretaan järjestelmään. Kiireellisissä tilanteissa radiologi saattaa joutua kirjoittamaan lausuntonsa käsin. B on lausuntoprosessi puheentunnistusta käyttäen. T_{Ima} on lausuntoprosessin mitattavissa oleva läpimenoaika.



tästä syystä puheentunnistukseen perustuva sanelumahdollisuus tulisikin olla käytettävissä jokaisessa radiologisessa työasemassa, jossa saneluja tehdään. Tällöin ei tarvitse erikseen odottaa vuoroaan puheentunnistustyöasemalle.

Puheentunnistuksen prosessin kulku

HUS-Röntgen on ollut mukana kehittämässä suomenkielistä radiologista puheentunnistusjärjestelmää, joka on ensimmäinen erikoissairaanhoidon tarkoitettu puheentunnistusjärjestelmä. Yhteistyöhanke käynnistyi talvella 2004 suomenkielisen radiologisen sanastoaineiston kokoamisella. Kootusta aineistosta luotiin ensimmäinen erikoissairaanhoidon suomenkielinen erikoissanasto ja puheentunnistusjärjestelmä (Philips SpeechMagic®, Philips Speech Recognition Systems GmbH). Erikoissanastoa kokeiltiin HUS-Röntgenin kahdessa röntgenyksikössä, ja keväällä 2005 pilotti laajennettiin Töölön sairaalan röntgenyksikköön.

Töölön röntgenyksikössä puheentunnistuskokeilu aloitettiin kolmen radiologin voimin, mutta sittemmin käyttäjien määrä on kasvanut. Nykyään kaikilla Töölön sairaalan röntgenyksikön radiologeilla on mahdollisuus käyttää puheentunnistusta, ja ohjelma on käytössä pilottiyksikön kaikissa radiologyöasemissa yhtä lukuun ottamatta.

Radiologisen lausuntoprosessin läpimenoaika (= aika siitä, kun tutkimus on suoritettu kuvantavalla laitteella, siihen, kunnes lausunto on kliinikon luettavissa) on mitattu useissa kansainvälisissä tutkimuksissa, erityisesti Yhdysvalloissa (7,8). Tutkimusten suora vertaaminen keskenään on hieman hankalaa, koska mitattavat prosessit eivät välttämättä ole täysin vertailukelpoisia keskenään. Uusimpien tutkimusten mukaan lausuntojen läpimenoaikojen lyheneminen on kuitenkin erittäin merkittävä (4,9).

Tutkimusryhmämme ensimmäisten kokemusten mukaan lausuntojen läpimenoaika lyheni noin puoleen (10) ja puheentunnistuksella tuotettujen lausuntojen osuus kaikista tehdyistä lausunnoista kasvaa kokemuksen karttuessa (11). Järjestelmän käytettävyyteen vaikuttaa myös kokemus siten, että hieman pidempään käyttäneiden käyttömäärä on suurempi kuin vasta-alkajien (11). Puheentunnistuksen osuus saneluiden tekemisessä kasvaa käyttökokemuksen myötä. Esimerkiksi alkuvuonna 2006 Töölön sairaalan röntgenyksikössä puheentunnistuksella tuotettujen saneluiden osuus kaikista 13 radiologin tuottamista saneluista oli 74 % ja loppuvuonna osuus oli 83 %.

Vertasimme niin ikään vuoden 2005 ensimmäisen vuosineljänneksen kasettipohjaisesti tehtyjen lausuntojen läpimenoaika vuoden 2006 ensimmäisen vuosineljänneksen puheentunnistuksella tehtyjen lausuntojen läpimenoaikaan. Otosväli oli molemmissa tutkimuksissa sama (1.1.–31.3.), mutta kasettipohjaista järjestelmää käytettäessä kiireellisissä tapauksissa radiologi joutui kirjoittamaan lausuntonsa itse. Käsinkirjoitettujen lausuntojen osuus kasettipohjaisista oli eri radiologeilla 10–20 %. Molemmissa otoksissa radiologit olivat samat ja he työskentelivät samassa yksikössä. Vertailuyksikön tutkimusprofiili ei merkittävästi muuttunut vertailun aikana ja tutkimuksia tehtiin lukumääräisesti lähes yhtä paljon.

Kasettipohjaisesti tehtyjä lausuntoja oli otoksessa 6 037 ja lausuntoprosessin läpimenoaika oli keskimäärin 24 h 46 min ± 76 h 31 min (keskiarvo ± SD). Puheentunnistusjärjestelmän avulla tehtyjä lausuntoja oli 6 486 ja lausuntoprosessin läpimenoaika oli 5h 23 min ± 27 h 42 min. Puheentunnistus nopeuttaa näin ollen lausuntoprosessia yli 78 % ($p < 0,0001$; riippumattomien otosten t-testi).

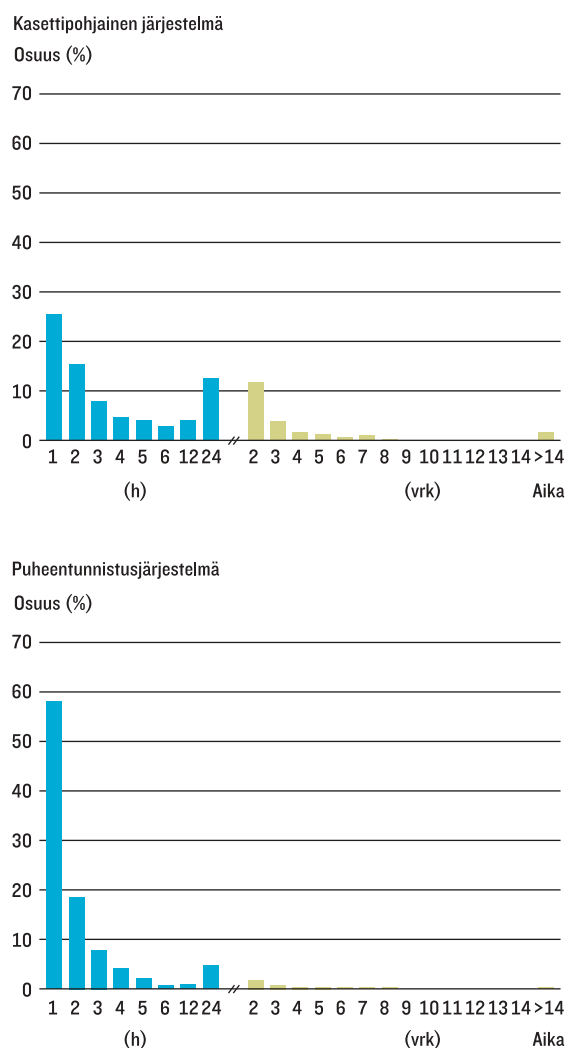
Puheentunnistuksella tuotetut lausunnot valmistuivat siis selvästi nopeammin kuin kasettipohjaisesti tuotetut (kuviot 2). Yli kolme neljäsosaa puheentunnistuksella tuotetuista lausunnoista oli valmiina kahden tunnin kuluessa.

Järjestelmän tuomat edut

Puheentunnistuksen tuoma ajansäästö on kiistaton. Alussa

Kuvio 2.

Lausuntojen läpimenoajan jakauma kasettipohjaista järjestelmää ($n = 6\ 037$; 1.1.–31.3.2005) ja puheentunnistusjärjestelmää ($n = 6\ 486$; 1.1.–31.3.2006) käytettäessä.



tosin radiologin yksittäiseen lausuntoon käyttämä aika voi hieman venyä, koska uuden järjestelmän käyttäminen vaatii harjaantumista ja itse puheentunnistusjärjestelmälle tulee luoda yksilöllinen ääniprofiili. Ääniprofiilin luomiseen ja koetekstien saneluun kannattaa kuitenkin käyttää aikaa, koska niillä pystyy parantamaan tunnistustarkkuutta. Lausuntojen korjauksiin tarvitaan varsin vähän aikaa, mutta yksittäisten virheellisten sanojen korjaamisella on suuri merkitys jatkoa ajatellen.

Integroitu radiologian tietojärjestelmä (RIS) ja kuvien arkistointijärjestelmä (picture archiving and communication system, PACS) helpottavat radiologin työtä, koska läheteet, kuvat ja lausunnot seuraavat toisiaan. Järjestelmien integrointi yhdistääkin sovellukset yhdeksi ratkaisuksi, jolloin niiden yhteiskäyttö helpottuu. Integroidussa järjestelmässä puheentunnistuslausuntoa aloitettaessa lausunto ohjautuu oikean potilaan läheteeseen ja kuvamateriaaliin. Puheentunnistusjärjestelmän käyttäminen on jopa vähentänyt lausuntojen joutumista väärän potilaan kohdalle ja helpottanut tässä suhteessa järjestelmävalvojen työtä. Puheentunnistus edellyttää tosin, että tutkimus on kirjattu ennen sanelun aloittamista ja että PACS-järjestelmään tallennetuissa tutkimuksissa on oikea tunnistusnumero.

Puheentunnistus on parantanut lausuntojen laatua ja rakennetta, koska radiologit pystyvät tarkistamaan välittömästi sanelutilanteessa sanojen oikeellisuuden, kun taas tekstinkäsittelijälle vaikeat termit ja ilmaisut tulee usein avata tai selittää saneluissa tarkemmin, jopa kirjain kirjaimelta tavaten. Perinteisessä konekirjoittajan purkamassa sanelussa syntyvät lapsukset jäävät kokonaan pois. Lisäksi kokemusten mukaan lausunnoista on tullut informatiivisempia, jäsentyneempiä ja suunnitelmallisempia, koska ne näkyvät ruudulla. Puheentunnistuksen myötä lausunnot ovat hieman lyhentyneet.

Merkittävin järjestelmän tuoma etu on kokonaisprosessin nopeutuminen ja se, että valmis lausunto on klinikon nähtävissä välittömästi lausunnon hyväksymisen jälkeen.

Lausuntojen tekemisessä voidaan hyödyntää myös puoli-valmiita lausuntopohjia ja helpottaa näin pitkien ja tyyppillisten lausuntojen tekemistä (esim. keuhko-TT tai vartalo-TT). Esimerkiksi pohja ”vartalon vakiolausunto 1” tuo näytölle peruslausunnon, jota radiologi voi täydentää ja korjata eikä hänen ei tarvitse tehdä koko lausuntoa alusta pitäen. Puheentunnistus mahdollistaa jatkossa myös strukturoitujen lausuntojen tekemisen, koska uudet puheentunnistusohjelmistojen versiot tukevat kansainvälisen standardin mukaista strukturoitua lausuntoa, johon puheentunnistus voidaan integroida (12,13,14,15). Täysin strukturoitujen lausuntojen tekoa tai lausuntopohjien käyttämistä ei vielä tässä vaiheessa kuitenkaan testattu.

Kliiniset kokemukset

Puheentunnistuksen käyttöönotto on helppoa ja vaatii lyhyen, keskimäärin 10–15 minuutin perehdytyksen. Jo muutaman ohjatusti tehdyn lausunnon jälkeen järjestelmää voi hyödyntää itsenäisesti. Uusien käyttäjien asetusten hie-

nosäädön ja perehdytyksen on suorittanut yksi HUS-Röntgenin radiologeista, jolloin järjestelmän käytössä eteen tulevat kysymykset ja ongelmat voidaan myös ratkaista nopeasti paikan päällä muiden konsultaatioiden lomassa. Oppimista nopeuttavat perustiedot sähköisestä radiologisesta lausuntolähetejärjestelmästä ja aikaisempi kokemus. Järjestelmän vastaanotto sekä sairaalan vakituisten radiologien että kuu-kausittain koulutuspestettä vaihtavien erikoistuvien sairaalalääkäreiden keskuudessa on ollut hyvä.

Puheentunnistusjärjestelmän integraation RIS- ja PACS-järjestelmiin tulee olla mahdollisimman saumatonta ja käytön yksinkertaista. Työasemalla olevan tutkimuksen lähete- ja sanelukenttä avautuvat automaattisesti ilman erillistä hakua. Tämä yhdistettynä puoliautomaatioituun tutkimusten hakuun (ns. wizardit) tekee sanelutyöstä varsin nopeaa. Tekstin muokkaamiseen voi esteettä ja vaihdellen käyttää sekä hiirtä, näppäimistöä että sanelumikrofonin painikkeita. Puheentunnistusta ohjaavana käyttöliittymänä (nauhoitus, kuuntelu, eteen- ja taaksekelaus) tavallisen sanelulaitteen liukukytintä jäljittelevä mikrofonin liukukytin on koettu luontevimmaksi, koska tämän jo selkäytimessä olevan hallintalaitteen käyttöä ei tarvitse erikseen opetella.

Tärkein hyöty Töölön sairaalassa on ollut lausuntojen nopeampi valmistuminen klinikon luettavaksi, mikä tukee nopean hoitoketjun toteutumista. Puheentunnistuksen nopeus voi korostua yksiköissä, joissa lausuntojen ei tarvitse olla valmiina niin nopeasti kuin traumasairaalassa.

Puheentunnistuksen tunnistustarkkuus on suhteellisen hyvä, mutta myös virheitä tulee ja lausunnon merkitys saattaa muuttua tai kääntyä jopa täysin päinvastaiseksi. Esimerkiksi ”luurakenne on skleroottista” voi vääristyä muotoon ”luurakenne on eroottista” ja ”luutumisen on edistynyt” muotoon ”luutumisen ei edistynyt”. Lausunnot on siis ehdottomasti tarkistettava huolellisesti ennen tallennusta. Tunnistus on osoittautunut toimivaksi, vaikka puhujan aksentti on melko vahvakin, kunhan käyttäjän suomen kielipin taidot muutoin ovat kohdallaan. Mikäli järjestelmään kuitenkin jatkuvasti syötettäisiin sanan väärää muotoa, sanan esiintymistilastot saattaisivat vääristyä ja lopulta sanan väärä muoto tulisi ensisijaiseksi. Käytännössä tällaista tilastollista ”köydenvetoa” on toistaiseksi ollut vain joidenkin kirjoitusasujen käytössä (discus ja diskus), niin että puheentunnistus tarjoaa ensisijaisesti enemmän käytettyä muotoa. Ahkerin sanelija on tässä demokratiassa voittaja.

Töölön sairaalan röntgenyksiköstä saatujen kokemusten mukaan käytännön radiologin työssä puheentunnistuksen hyödyt ylittävät kuitenkin selvästi haitta-vaikutukset.

Kirjallisuutta

- 1 White KS. Speech recognition implementation in radiology. *Pediatr Radiol* 2005;35:841-6.
- 2 Vorbeck F, Ba-Ssalamah A, Kettenbach J, Huebsch P. Report generation using speech recognition in radiology. *Eur Radiol* 2000;10:1976-82.
- 3 Virtanen J. Lääkäri puhuu – kone kirjoittaa. *Suom Lääkäril* 2005;60:3896.
- 4 Trumm C, Francke M, Küttner B, Nissen-Meyer S, Reiser M, Glaser C. Speech recognition: impact on report availability and clinical workflow. *Hospital Imaging&Radiology Europe* 2006;1:1-6.
- 5 Deng L, Erler K. Structural design of hidden Markov model speech recognizer using multivalued phonetic features: comparison with segmental speech units. *J Acoust Soc Am* 1992;92:3058-67.
- 6 <http://www.stm.fi> ja <http://www.terveysshanke.fi>
- 7 Langer SG. Impact of speech recognition on radiologist productivity. *J Digit Imaging* 2002;15:203-9.
- 8 Langer SG. Impact of tightly coupled PACS/speech recognition on report turnaround time in the radiology department. *J Digit Imaging* 2002;15:234-6.
- 9 Rana DS, Hurst G, Shepstone L, Pilling J, Cockburn J, Crawford M. Voice recognition for radiology reporting: Is it good enough? *Clin Radiol* 2005;60:1205-12.
- 10 Kauppinen T, Tolkkio O, Korhonen P, Ahovuo J. Using speech recognition reduces time slot between imaging and final report. *Eur Radiol* 2005;15 suppl I:182.
- 11 Kauppinen T, Ahovuo J. Usability of speech recognition in a trauma hospital: One year follow-up study. *Eur Radiol* 2006;16 suppl I:166.
- 12 Liu D, Zucherman M, Tulloss Jr WB. Six characteristics of effective structured reporting and the inevitable integration with speech recognition. *J Digit Imaging* 2006;19:98-104.
- 13 Reiner B, Siegel E. Radiology reporting: Returning to our image-centric roots. *AJR* 2006;187:1151-5.
- 14 Talton D. Perspectives of speech recognition technology. *Radiol Manage* 2005;27:38-40, 42-43.
- 15 Eng J, Eisner JM. Radiology report entry with automatic phrase completion driven by language modeling. *Radiographics* 2004;24:1493-501.

Tomi Kauppinen

dosentti, teknologiapäällikkö
HUS-Röntgen
tomi.kauppinen@hus.fi

Mika Koivikko

IT, erikoislääkäri
HUS-Röntgen

Seppo Koskinen

dosentti, osastonylilääkäri
HUS-Röntgen

Juhani Ahovuo

dosentti, toimitusjohtaja
HUS-Röntgen

English summary

Finnish speech recognition in radiology

Speech recognition has been used in medicine for dictating reports, especially in radiology, in many pan-European languages. A separate special vocabulary (so-called context) and language model is required for each medical speciality, because context increases recognition level significantly. The first Finnish speech recognition system in specialised medical care was developed for radiology.

In this study we tested and piloted Finnish radiological speech recognition in Töölö trauma hospital in the department of radiology. Our finding was that speech recognition decreased turnaround times by more than 78%, when cassette based and speech recognition processes were compared. Speech recognition enables clinicians to read reports immediately after the radiologist has completed them. It was discovered that report turnaround times decreased from 25 h to 5.5 h. This difference was statistically significant ($p < 0.0001$). Learning to use the speech recognition system was fairly easy for new users. An orientation period of only 10 to 15 minutes was required before making the first dictations.

It was also discovered that speech recognition improves the informative value, structure and clarity of the radiological reports. In conclusion, speech recognition speeds up and improves the quality of the patient care process.

Tomi Kauppinen

Ph.D., Technology Manager, HUS Helsinki Medical Imaging Center
Hospital District of Helsinki and Uusimaa
E-mail: tomi.kauppinen@hus.fi

Mika Koivikko

Seppo Koskinen
Juhani Ahovuo