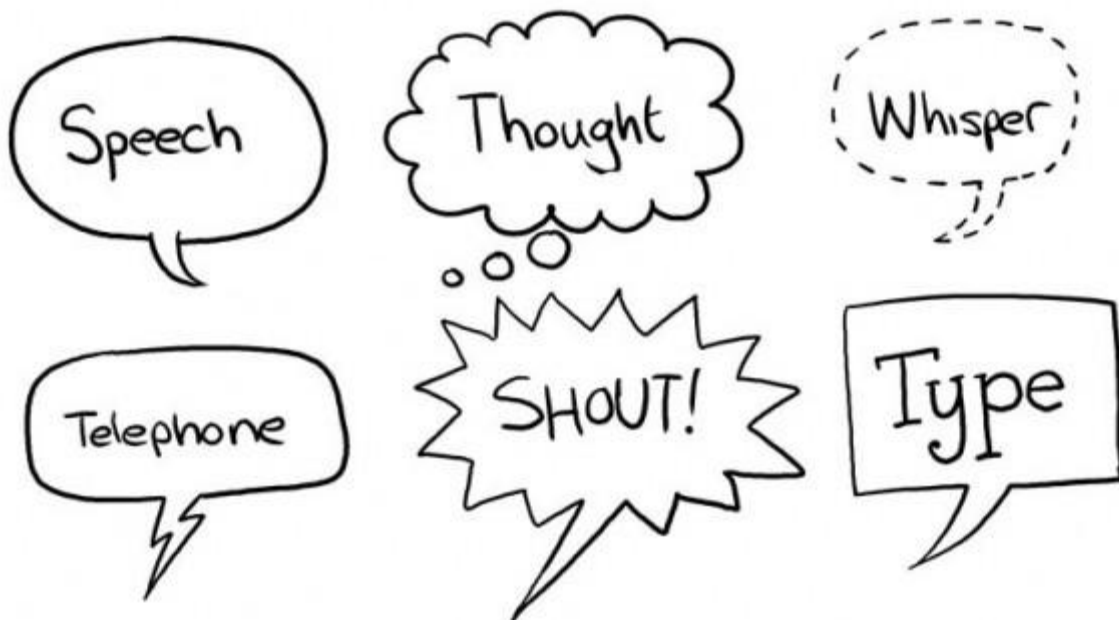


Bacheloropdracht

*Spraakherkenning binnen een bedrijfsomgeving:
een oplossing voor e-mail overload?*



Joost Gooskens
Bruggertstraat 325
7545 AV Enschede
Studentnummer: 9602089

Begeleiders:
IR. H. Kroon
Dr. P.C. Schuur

Management Summary

Vanaf begin jaren 80, wanneer er sprake is van een explosieve groei van het gebruik van e-mail, zijn er de eerste signalen dat het gebruik van dit communicatiemedium ook negatieve effecten met zich meedraagt. Men ontvangt meer en meer e-mail en een steeds groter gedeelte daarvan is ongewenst. Begin jaren 90 komt voor het eerst de term *information overload* voor in de literatuur. Vlak daarna is er ook sprake van *e-mail overload*, te weten het gevoel niet meer om te kunnen gaan met het aanbod aan informatie die via e-mail binnenkomt. Dit neemt dermate grote vormen aan dat in 2004 door de editor-in-Chief van de MIS Quarterly een editorial aan wordt gewijd. Hij vraagt zich af of het gebruik van *voice recognition software* (VRS) bij kan dragen aan het doen afnemen van de overloadgevoelens die met de afhandeling van e-mail samenhangen.

In dit onderzoek is door middel van een literatuurstudie gekeken of hier bewijs van te vinden is in wetenschappelijke publicaties in de belangrijkste journals op het gebied van de Information Technology. Na bestudering van 68 publicaties in verschillende journals is gebleken dat er geen duidelijke relatie te leggen is tussen het gebruik van VRS en de afname van overloadgevoelens. Er is zelfs geen bewijs te vinden dat er gebruik gemaakt wordt van deze technologie bij de afhandeling van e-mail in een bedrijfsomgeving. Vervolgens is er gekeken naar de manier waarop bedrijven omgaan met VRS, of deze technologie in bedrijven wel wordt gebruikt en wat eventueel de toekomstperspectieven zijn met betrekking tot het gebruik van VRS.

Het blijkt dat door experts en wetenschappers uit verschillende disciplines binnen de IT het potentieel voor het gebruik van VRS wordt onderkend. De technologie is begin 21^e eeuw dermate ver ontwikkeld dat een herkenningssnauwkeurigheid van 95% haalbaar is met de nu beschikbare softwarepakketten. De software is in de nieuwste versies bijzonder gebruiksvriendelijk en er zijn pakketten beschikbaar voor specialistische bedrijfstakken. Toch is er vanuit de literatuur nauwelijks bewijs te vinden dat VRS op grote schaal binnen het bedrijfsleven wordt gebruikt.

Een positieve uitzondering hierop is het veld van de medische radiologie. Binnen deze bedrijfstak worden erg veel rapporten door de specialisten gegenereerd. Dit gebeurt nog vaak door traditionele transcripties, maar het gebruik van VRS is hier de laatste jaren flink in opkomst. Een nadeel dat in de praktijk naar voren blijkt te komen is dat de nauwkeurigheid van VRS toch te wensen overlaat. In rapporten gegenereerd met VRS zitten nog veel fouten, en sommige fouten kunnen potentieel grote gevolgen hebben voor de zorg die verleend wordt aan patiënten. Het is dus duidelijk dat er nog goed gecontroleerd moet worden of rapporten die met VRS tot stand gekomen zijn wel accuraat genoeg zijn.

Het voordeel dat behaald wordt met het gebruik van VRS wordt door deze benodigde extra controle tenietgedaan. De tijdsbesparing en hiermee samenhangende verwachte kostenbesparing zijn voor ziekenhuizen vaak de reden om gebruik te gaan maken van VRS. In de praktijk blijkt dat de specialisten veel van hun kostbare tijd kwijt zijn met de verbetering van fouten en het is dus maar de vraag of de positieve effecten opwegen tegen de negatieve effecten. Wel lijkt de radiologie te kunnen leven met de tekortkomingen van VRS en veel instellingen stappen over op het gebruik van VRS ten koste van de traditionele transcriptiemethodes.

Of in de toekomst meer bedrijfstakken gebruik zullen gaan maken van VRS valt nog te bezien. Feit is dat de technologie zich nog altijd verder ontwikkelt en dat mensen zich steeds makkelijker aanpassen aan technologische ontwikkelingen. Wellicht is er toch nog een gouden toekomst weggelegd voor het gebruik van VRS.

Voorwoord

Voor u ligt het onderzoek waarmee ik mijn bachelor van de studie Technische Bedrijfskunde aan de Universiteit Twente afsluit. Het is een reis met hoogtenpunten, dalen, veel plezier, studie-ontwijkend gedrag, bestuursfuncties en momenten van tanende motivatie geweest. Een geweldige tijd op mijn studentenflat, de Schildpatio, heeft mij een onvergetelijke, doch zeer lange studententijd bezorgd. Mijn ouders hebben af en toe met hun handen in het haar gezeten en zullen zich regelmatig afgevraagd hebben of dit moment ooit zou plaatsvinden. Het is me toch gelukt!

De opdracht hangt al twee jaar boven mijn hoofd. Na opstartproblemen en het vertrek uit Nederland van mijn oorspronkelijke begeleider mocht ik onder begeleiding van Henk Kroon toch mijn onderzoek afronden. Mijn inzichten over de onderzochte technologie, die al jaren bekend staat als hoopgevend en veelbelovend, zijn flink verbeterd. De kloof tussen de resultaten van wetenschappelijke onderzoek en het werkelijke gebruik van technologie binnen het bedrijfsleven hebben mij in eerste instantie verbaasd, maar ik begrijp uiteindelijk de redenen die ten grondslag liggen aan het gebruik van nieuwe technologieën als spraakherkenning binnen een (commerciële) bedrijfsomgeving.

Een belangrijke vriend, vooral de laatste jaren, is ook Mark Tangeman geweest. Als mede langstudeerder hebben we elkaar proberen te motiveren en via hem ben ik uiteindelijk bij Henk Kroon, mijn afstudeerbegeleider, beland. Henk wil ik ook zeker graag bedanken voor het geduld dat hij met mij heeft gehad. Ook wil ik Dr. Schuur bedanken voor het optreden als mijn tweede begeleider. Voor het feit dat mijn beide begeleiders, ondanks de vakantieperiode, wilden meewerken aan een spoedig afstuderen ben ik ze zeer erkentelijk.

Tenslotte wil ik mijn grote liefde, steun en toeverlaat en toekomstige vrouw Anne-Margreeth bedanken. Via de bemoeizucht van twee goede vrienden hebben we elkaar tijdens een voetbalwedstrijd van het Nederlands elftal ontmoet. Sindsdien is ze mijn grootste motivator geweest, alhoewel dit bij tijd en wijle voor haar erg moeilijk moet zijn geweest. De schop onder mijn kont die ik nodig had heeft ze vaak, misschien iets té vaak, uit moeten delen. Gelukkig heeft dit het gewenste resultaat gehad en daarom ligt dit verslag nu voor uw neus. Lieverd, bedankt! Ik hou van je!

Joost Gooskens

Enschede, 21 augustus 2011

Inhoudsopgave

Management Summary	2
Voorwoord	3
Inhoudsopgave	4
1 Projectbeschrijving	5
1.1 Achtergrond	5
1.2 Oorzaak.....	5
2 Literatuuronderzoek.....	6
2.1 Doel en Onderzoeksvragen.....	6
2.2 Onderzoeksbeschrijving	6
2.3 Resultaten onderzoek	6
3 Beschrijving spraakherkenning	9
3.1 Definitie spraakherkenning.....	9
3.2 Definitie (e-mail) overload.	9
4 Huidige stand van zaken (d.d. februari 2011)	11
4.1 Inleiding	11
4.2 Afbakening literatuur.....	11
4.3 Tijdelijk elektronische communicatie en overload.....	11
4.4 Tijdelijk Spraakherkenning	15
5 Antwoord op de onderzoeksvragen	27
5.1 Wetenschappelijke visie	27
5.2 VRS en e-mail	27
5.3 Toekomstperspectief VRS.....	28
6 Conclusie.....	30
Referenties.....	31

1 Projectbeschrijving

In dit hoofdstuk wordt uitgelegd waarom een literatuuronderzoek naar het gebruik van spraakherkenning in bedrijfsomgevingen, voornamelijk bij de afhandeling van e-mail, is uitgevoerd.

1.1 Achtergrond

E-mail is een dienst die in de moderne tijd, en dus ook in het dagelijkse bedrijfsleven, niet meer weggedacht kan worden. Wereldwijd zijn er vele gebruikers (volgens een onderzoek van de Radicati Groep in 2008 1,3 miljard unieke gebruikers ^{a)}) en ook binnen bedrijfsomgevingen wordt e-mail als belangrijk communicatiemiddel gezien. Volgens dezelfde Radicati Group waren er eind 2007 516 miljoen zakelijke inboxen actief. Geografische grenzen zijn met dit communicatiemiddel niet belangrijk meer en het tijdsaspect vervaagt ook steeds meer. Er kleven echter ook nadelen aan het gebruik van e-mail op deze schaal. Mensen verwachten tegenwoordig snel een antwoord op verzoeken en vragen die ingediend worden via de e-mail. Het neemt de functie van de telefoon gedeeltelijk over, en dit zorgt ervoor dat mensen ook voor het ontvangen en versturen van e-mail continu bereikbaar dienen te zijn. Dit zorgt ook voor een grote inbreuk op het privéleven van mensen.

E-mail overload¹ zorgt ook voor een steeds grotere druk bij managers op de werkvloer. Het gebruik van e-mail binnen bedrijven heeft de laatste jaren een grote vlucht genomen. Doordat een manager tegenwoordig dagelijks gemiddeld tientallen e-mails binnen kan krijgen, die op zeer korte termijn beantwoord dienen te worden, neemt het gevoel van overload toe. E-mails blijven onbeantwoord en inboxen lopen over van in behandeling zijnde e-mail. Dit heeft zijn weerslag op de tijd die managers hebben om hun overige taken uit te voeren. Deze tendens heeft Ron Weber, Editor-in-Chief van het toonaangevende journal MIS Quarterly, ertoe bewogen om in 2004 een editorial² te schrijven die dit probleem adresseert. Hij detecteert een achterblijvend onderzoek naar het gebruik van e-mail in organisaties en draagt enkele (ongefundeerde) oplossingen aan om het gevoel van overload te verminderen. Een van de oplossingen die hij als erg hoopgevend ziet is het gebruik van spraakherkenningssoftware.

1.2 Oorzaak

Naar aanleiding van de problemen die samenhangen met e-mail overload wordt er aan de universiteit Twente een onderzoek uitgevoerd naar e-mail overload. Dit onderzoek probeert in kaart te brengen in hoeverre e-mail overload ook echt een probleem vormt bij bedrijven, welk effect het heeft op het functioneren van mensen, wat de belangrijkste oorzaken zijn voor het gevoel van overload en tenslotte probeert men ook enige oplossingsrichtingen te onderzoeken. Dit bachelorverslag hoopt bij te dragen aan de laatste fase van dit onderzoek.

Dit onderzoek probeert via een literatuurstudie uit te vinden of deze oplossing ook echt kans van slagen kan hebben. Ook is een van de doelen om inzicht te verkrijgen in het huidige gebruik van spraakherkenning. Wordt er al gebruikt van gemaakt door bedrijven. Zo ja, hoe functioneert het? Zo nee, waarom niet?

Tenslotte kan dit onderzoek ook de basis vormen voor vervolgonderzoek. In ieder geval zal het een goed overzicht geven van wat er in de literatuur bekend is over dit onderwerp.

2 Literatuuronderzoek

In het komende hoofdstuk zal uitgelegd worden hoe het literatuuronderzoek precies is uitgevoerd. Eerst zal er een overzicht gegeven worden van het onderzoeksdoel en de onderzoeksvragen. Vervolgens zal er uitgelegd worden welke onderzoeksstrategie er gevolgd is, welke databases zijn gebruikt en waarom en op basis waarvan de selectie uit de gevonden artikelen is gemaakt. Tenslotte zal er een overzicht gegeven worden van alle relevante artikelen die bij dit onderzoek naar boven zijn gekomen.

2.1 Doel en Onderzoeksvragen

Onderzoeksdoel:

Het doel van dit onderzoek is het maken van een literatuursynthese over het gebruik van spraakherkenning in bedrijfsomgevingen. Door een grondig literatuuronderzoek uit te voeren kan er een goed inzicht verkregen worden in de literatuur die tot op heden verschenen is over dit onderwerp. Ook is er gezocht naar literatuur over het gebruik van spraakherkenning in een e-mail omgeving in het bijzonder.

Onderzoeksvragen

1. Wat is binnen de wetenschappelijke publicaties de visie met betrekking tot het gebruik van spraakherkenningssoftware in een bedrijfsomgeving waar veel gebruik gemaakt wordt van e-mail?
2. Zou het gebruik van spraakherkenningssoftware de afhandeling van e-mail kunnen versnellen?
3. Heeft het gebruik van spraakherkenningssoftware een toekomst binnen een bedrijfsomgeving?

2.2 Onderzoeksbeschrijving

Dit onderzoek is uitgevoerd in een aantal stappen. De eerste stap betreft het opstellen van het onderzoeksdoel en de onderzoeksvragen. Dit is in de voorgaande paragraaf al beschreven.

Stap 2 betrof het bepalen van de verschillende informatiebronnen die geraadpleegd worden.

De derde stap betrof het zoeken van relevante artikelen. Door gebruik te maken van slim gekozen keywords is er gezocht naar publicaties die van belang kunnen zijn voor deze literatuurstudie. Deze keywords zullen later in dit verslag nog beschreven worden.

De volgende stap betrof de verfijning van de gevonden resultaten. Er is gekeken naar relevantie, wetenschappelijk niveau van de publicatie en bruikbaarheid bij beantwoording van de onderzoeksvragen. Hierdoor viel een gedeelte van de artikelen af. De overgebleven artikelen zijn vervolgens volledig doorgelezen.

Als laatste is er nog gekeken naar de literatuurverwijzingen in deze artikelen. Mochten de artikelen waarnaar verwezen werd niet eerder naar boven zijn gekomen via de zoekopdrachten dan zijn deze alsnog geprobeerd te vinden. Deze stap staat ook wel bekend als de 'backward reference search'.

De artikelen die na deze stappen overgebleven zijn vormen de basis van de literatuursynthese. In hoofdstuk vier zal worden samengevat wat de belangrijkste conclusies waren met betrekking tot het onderwerp van dit onderzoek.

2.3 Resultaten onderzoek

Stap 1 zal in deze paragraaf niet verder toegelicht worden. Dit aangezien in paragraaf 2.1 het onderzoeksdoel en de onderzoeksvragen al besproken zijn.

Bij stap 2 is er een keuze gemaakt uit de beschikbare informatiebronnen. In dit onderzoek betreft het uitsluitend digitale informatiebronnen. Er is getracht om alle belangrijke journals binnen het gebied van de IS (Information Systems) te bestrijken. Hierbij is gebruik gemaakt van het onderzoek uitgevoerd door Schwartz en Russo¹, waarin beschreven wordt welke zoekmachines het beste gebruikt kunnen worden om een zo groot mogelijke dekking van de top 25 van de IS Journals te bereiken. Er is voor gekozen om gebruik te maken van de databases waarop de Universiteit Twente geabonneerd is (Web Of Science, PiCarta, Scopus), uitgebreid met de databases van ACM en Ingenta, welke bij het onderzoek van Schwartz en Russo als databases met de beste dekking van de Top 25 naar voren kwamen. Mochten de publicaties gevonden zijn een journal van een ander veld dan IS dan zijn deze artikelen, indien relevant, ook meegenomen in dit onderzoek.

In stap 3 zijn 83 artikelen gevonden. Er is gebruik gemaakt van de zoekmachines uit stap 2 en er is gezocht met de volgende keywords:

- Speech AND/OR Voice AND/OR Recognition AND/OR Information Systems
- E-mail AND Overload

De volgende stap in het onderzoek is de analyse van deze artikelen geweest. Van alle artikelen zijn de managementsamenvattingen (indien aanwezig) doorgenomen en tevens de conclusies. Naar aanleiding van de informatie die hieruit naar voren kwam zijn een aantal artikelen afgevallen omdat ze niet-relevante informatie bevatten, of ingingen op onderwerpen die buiten de scope van dit onderzoek vallen. Hierdoor bleven 75 artikelen over die relevante informatie met betrekking tot de onderzoeksvragen bevatten. De volgende inclusie- en exclusiecriteria zijn bij deze selectie gebruikt:

Inclusiecriteria

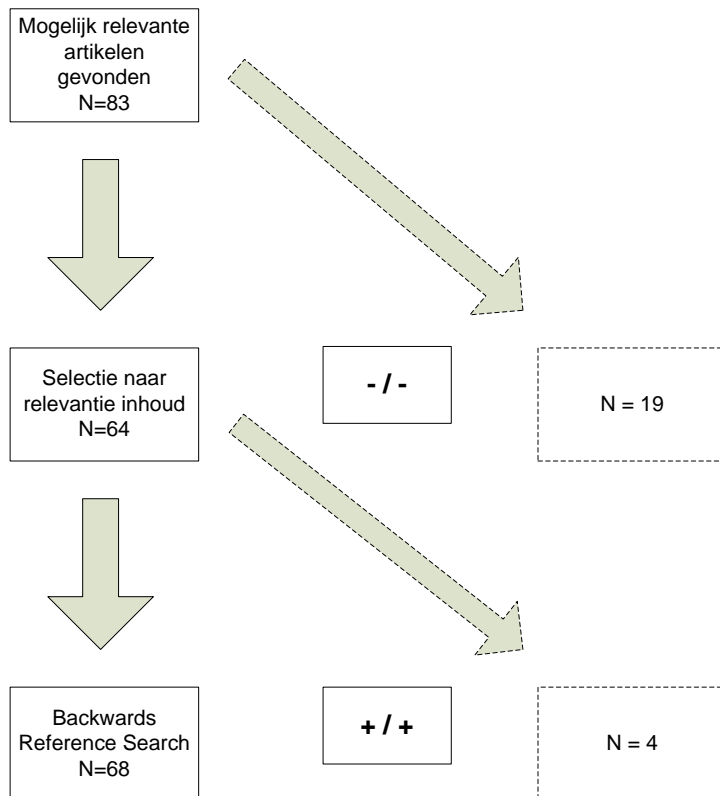
- Het betreft artikelen die het gebruik van spraakherkenning binnen een bedrijfsomgeving beschrijven
- Het betreft artikelen die de procedures over e-mailgebruik behandelen waarbij op enigerlei manier gebruikt wordt van spraakherkenningssoftware.
- Het betreft artikelen die de huidige staat van de techniek van spraakherkenning beschrijven en/of uitspraken doen over ontwikkelingen in de nabije toekomst.
- Het betreft artikelen die het verband tussen het gebruik van spraakherkenningssoftware en daar eventueel mee samenhangende tijdsbesparingen beschrijven.

Exclusiecriteria

- De artikelen behandelen technieken die inmiddels gedateerd zijn
- De artikelen voldoen niet aan de kwaliteitseisen. Dit kan blijken uit het feit dat ze wel te vinden zijn op het web, maar niet officieel gepubliceerd zijn.
- Het betreft artikelen die spraakherkenning gebruiken op het 'woord-niveau' en niet op 'spraak-niveau'. Dit kan gezien worden als de herkenning van complete zinnen ten opzichte van herkenning van losse namen en/of woorden. Een uitgebreidere beschrijving komt aan bod in hoofdstuk 3.

De volgende en laatste stap was het uitvoeren van een 'backwards reference search'. Er is gekeken naar de literatuurverwijzingen uit de relevante artikelen. Naar aanleiding van deze references zijn er nog 10 artikelen toegevoegd aan de lijst met relevante publicaties. Het totaal aan relevante artikelen is zo dus uitgekomen op 85 en deze artikelen zijn vervolgens volledig doorgelezen. De belangrijkste speerpunten en conclusies zijn samengevat in hoofdstuk vier.

De gebruikte zoek- en selectieprocedures zijn samengevat in figuur 1. Hier is ook te zien hoeveel artikelen in elke stap afgevallen zijn en hoeveel er lopende het onderzoek eventueel toegevoegd zijn.



Figuur 1: selectieprocedure artikelen

Via een op chronologie gebaseerde beschrijving is in hoofdstuk 4 een samenvatting gegeven van de opvattingen die door de jaren heen uit gepubliceerde wetenschappelijke onderzoeken naar boven zijn gekomen.

In hoofdstuk 5 zullen vervolgens de onderzoeksvragen zoals beschreven in paragraaf 2.1 beantwoord worden. Hierbij zal gebruik gemaakt worden van de inzichten die verkregen zijn na bestudering van de literatuur.

Tenslotte zullen de belangrijkste conclusies gegeven worden in het laatste hoofdstuk. Hierbij zal ook de opinie van de auteur gegeven worden naar aanleiding van het onderzoek en ook zal er een korte toekomstvisie gegeven worden. Of deze visie in de toekomst werkelijkheid zal worden zal de komende jaren moeten blijken.

3 Beschrijving spraakherkenning

In dit hoofdstuk zal de theoretische achtergrond van dit onderzoek kort behandeld worden. Er zullen een aantal definities gegeven worden en enkele begrippen zullen kort uitgelegd worden.

3.1 Definitie spraakherkenning

Het spraakherkenningsproces, uitgevoerd door computers, doorloopt over het algemeen dezelfde stappen. Deze stappen zullen nu kort toegelicht worden.

1. Geluidsinvoer. Er wordt gesproken in een microfoon, welke aangesloten is op een computer met een geluidskaart.
2. Het geluid wordt verwerkt door de geluidskaart van de computer. Dit gebeurt over het algemeen door een akoestische processor, welke het achtergrondgeluid eruit filtert en het geluid opdeelt in een serie van fonema. Een foneem is een verzameling van klanken die allemaal dezelfde betekenisonderscheidende functie hebben. Dit zijn de kleinste bouwstenen van spraak.
3. Woordvergelijking. De software probeert dit op twee manieren te doen. Als eerste probeert het met een akoestische analyse een lijst samen te stellen met mogelijke matchende woorden die gelijke klanken of fonema bevatten. Vervolgens gebruikt de software 'language modeling' (de aannemelijkheid dat een woord komt tussen het voorgaande en opvolgende woord) om deze lijst te verkorten. Hier kan ook gebruik gemaakt worden van een specifiek woorddomein dat door de gebruiker gedefinieerd is. Dit domein kan samengesteld zijn uit de training van de gebruiker met het herkenningssysteem op het gebied van woordkeuze, uitspraak en dergelijke, als wel met standaardpakketten voor bijvoorbeeld medische, wettelijke en/of overige toepassingen die speciaal voor de spraakherkenningssoftware ontwikkeld zijn. De continue-spraak-software interpreteert ook contextuele informatie om te kunnen voorspellen welke woorden de grootste kans hebben om in de vervolgzin gebruikt te worden. Dit helpt het systeem ook om onderscheid te maken tussen homoniemen (verschillende woorden die door de klanken erg op elkaar lijken, zoals bijvoorbeeld dank en bank). De gedetecteerde woorden worden tenslotte in de meest logische zinsconstructie geplaatst.
4. Tekst-output. Sommige software heeft zijn eigen ingebouwde wordprocessor, maar vaak is het mogelijk om een gangbaar wordprocessing programma (MS Word, Excel, wordpad etc.) aan de software te koppelen. Zelfs een tekstbox in andere applicaties, zoals webbrowsers en e-mailcliënten kan hiervoor gebruikt worden.

De meest gebruikte spraakherkenningssoftware die momenteel beschikbaar is bestaat uit CMU Sphinx (Open source), MacSpeech (voor Macintosh, gebaseerd op Philips Speech Technology, Windows Speech Recognition (ingebouwd in de Windows Vista en Windows 7 platforms), Sonic Extractor (vooral voor telefonie en omroep-toepassingen), Dragon NaturallySpeaking en Tatzi (bron: wikipedia).

Uit de literatuur blijkt dat Dragon NaturallySpeaking het pakket is dat het meest gebruikt wordt voor onderzoeken en commerciële toepassingen. Binnen de medische wereld, en dan vooral het terrein van de radiologie, wordt in Europa het meest gebruik gemaakt van Philips SpeechMagic of hiervan afgeleide software.

3.2 Definitie (e-mail) overload.

Het begrip e-mail overload is een onderdeel van het overkoepelende begrip 'information overload'. Ter verduidelijking worden beide begrippen in deze paragraaf kort uitgelegd.

Information overload wil zeggen dat de hoeveelheid informatie die ontvangen wordt groter is dan de capaciteit van de ontvangende persoon om deze informatie te verwerken⁴.

Als we hier wat verder op inzoomen kunnen we een definitie geven van het begrip e-mail overload. Dit bekijkt het overload principe vanuit het oogpunt van de ontvangen e-mail. E-mail overload kan dan omschreven worden als information overload die voortvloeit uit a) de hoeveelheid berichten die elke dag verstuurd en ontvangen worden, b) de manier waarop deze berichten geschreven worden (taalgebruik, gebruik van cc en bcc, netiquette-regels die overtreden worden etcetera) en c) de verschillende manieren waarop werknemers gebruik maken van hun e-mailsysteem.

Volgens het artikel "**Reconceptualizing e-mail overload**"⁵ zijn de 5 belangrijkste veroorzakers van e-mail overload: 1) e-mail is erg gemakkelijk te distribueren, 2) gebruikers staan altijd in contact met elkaar, 3) gebruikers kunnen gemakkelijk multitasken op manieren die voorheen niet bestonden, 4) taken zijn extreem complex en 5) men verwacht onmiddellijk een antwoord.

Als de hoeveelheid van ontvangen (en dus te behandelen) mail dermate toeneemt dat de ontvanger last krijgt van stressgevoelens spreken we dus van e-mail overload. E-mail wordt in de moderne bedrijfsomgeving gezien als de belangrijkste communicatievorm en een relatief simpel mailtje van 4 regels kan een verzoek of opdracht bevatten die een enorme druk kan leggen op de tijdsbesteding van een medewerker. Ook het feit dat men e-mail tegenwoordig ziet als vervanger van de telefoon en andere synchrone communicatiemiddelen en dus eigenlijk op erg korte termijn antwoord verwacht op verstuurd berichten draagt bij aan de verhoging van het gevoel van stress die e-mail kan veroorzaken.

4 Huidige stand van zaken (d.d. februari 2011)

4.1 Inleiding

In de literatuur zijn veel voorbeelden te vinden van het gebruik van spraakherkenningssoftware. De techniek heeft de laatste jaren een grote vlucht genomen. Dit heeft vooral te maken met de ontwikkelingen op het gebied van de IT in het algemeen. De rekencapaciteit van computers neemt elk jaar toe. Tevens is de personal computer niet meer weg te denken uit de maatschappij. Ook het bedrijfsleven heeft de automatisering omarmd en er zijn bijna geen bedrijven actief die in meer of mindere mate gebruik maken van computers.

Voor de term Voice Recognition Software zal in het vervolg van dit verslag de afkorting VRS gebruikt worden.

In de eerste paragraaf zal beschreven worden hoe de onderzochte literatuur is afgebakend. De VRS technologie wordt op verschillende niveaus gebruikt en niet al het gebruik valt binnen de scope van dit onderzoek. In de paragrafen 4.3 en 4.4 zal via een tijdlijn, dus op chronologische volgorde, worden beschreven wat binnen de wetenschappelijke literatuur de belangrijkste conclusies zijn met betrekking tot elektronische communicatie en overload en het gebruik van spraakherkenning in het algemeen en VRS in het bijzonder binnen bedrijfsomgevingen.

4.2 Afbakening literatuur

Binnen de telecomwereld wordt ook steeds meer gebruik gemaakt van spraakherkenning. Op de moderne mobiele telefoons, de zogenaamde smartphones, zit ook vaak spraakherkenning. Het betreft echter vooral functionaliteit om het kiezen van telefoonnummers uit het telefoonboek te versnellen. De uitgesproken naam wordt vergeleken met een vooraf opgenomen audiospoor en als het goed is herkend. Daar deze technologie verre van foutloos werkt en er dus vaak meerdere malen gesproken moet worden voordat de goede naam herkend wordt gebruiken telefoonbezitters deze functionaliteit niet veel. Er is wel een ontwikkeling gaande om ook spraakherkenning te gaan gebruiken bij het opstellen van SMS-berichtjes, maar hier is in de literatuur nog niets over te vinden. Aangezien het gebruik van spraakherkenning bij het gebruik van mobiele technologie in het algemeen buiten de scope van dit onderzoek valt zal hier verder niet op ingegaan worden.

Ook telefooncentrales maken veelvuldig gebruik van spraakherkenningstechnologie. Als men vandaag de dag contact opneemt met een service- of informatienummer is de kans groot dat men eerst een spraakgestuurd menu moet doorlopen voordat men een fysiek persoon aan de lijn krijgt. Het spraakherkenningsgedeelte bestaat vaak uit herkenning van klantnummers, plaatsnamen of storingsmeldingen. Omdat deze technologie vaak niet naar behoren werkt wekken dit soort menu's vaak irritatie op bij bellende klanten. Wel is er een verbetering van de herkenning in gang gezet die veelbelovend is voor de toekomst. Dit gebruik van spraakherkenning valt echter ook buiten de scope van dit onderzoek en zal ook verder niet behandeld worden.

4.3 Tijdlijn elektronische communicatie en overload

Al in 1982 wordt door Peter Denning, president van ACM, gesignaleerd dat de ontwikkeling in het gebruik van e-mail ook minder positieve neveneffecten heeft. De opkomst van junkmail zorgt van een toenemende druk op ontvangers. Hij vraagt zich dan al af of er ooit een bescherming komt voor de ontvangers voordat zij 'verdrinken' in de zee van informatie die gegenereerd gaat worden⁶.

In 1985 blijkt uit een publicatie van Starr Roxanne Hiltz, een inmiddels gepensioneerde professor werkzaam op de afdeling Information Science/Information Systems op het New Jersey Institute of Technology (NJIT) en in samenwerking met haar man Murray Turoff auteur van "The Network Nation, een boek dat beschreven wordt als een van de

eerste boeken die de elektronische grens afbakent, dat overload kan ontstaan bij de afhandeling van mail. Hun conclusie is dat software een klein gedeelte van de oplossing kan zijn, maar dat het grootste gewin te halen is uit het veranderen van het gedrag van de gebruikers⁷.

In 1995 signaleert Janet Fulk dat het gebruik van elektronische communicatie het ontwerp van de organisatie flink zal beïnvloeden. Er moet een betere koppeling komen tussen de gebruikte technologiesystemen en het organisatiesysteem. De configuratie van de mailprotocollen binnen een bedrijf moeten zorgen voor een soepelere afhandeling van e-mail en zij concludeert dat er betere theorieën en methodes ontworpen moeten worden voor de afhandeling van (elektronische) asynchrone communicatie en dat in de toekomst waarschijnlijk adaptieve systemen ontwikkeld zullen worden⁸.

De noodzaak hiervan blijkt ook uit het rapport 'Dying for Information', opgesteld door Reuters en beschreven door Charles Arthur, science editor van The Independent, waarin geconcludeerd wordt dat de helft van de 1300 ondervraagde managers klagen over information overload. Een groot gedeelte hiervan wordt veroorzaakt door de enorme hoeveelheid, vaak ongevraagde, informatie die ze toegestuurd kregen. Hij legt voor het eerst het grootste gedeelte van de schuld bij het interne gebruik van e-mail. De groep ondervraagde managers voorzag ook dat het internet de belangrijkste oorzaak van dit probleem zou vormen in de jaren erna⁹.

Steve Whittaker³ concludeert uit zijn onderzoek dat het gebruik van e-mail ook aan verandering onderhevig is. Was het eerst een pure communicatievorm, tegenwoordig gebruiken mensen hun inbox onder andere als 'to-do-lijst' waarin ongelezen berichten aangeven dat er nog actie ondernomen dient te worden, als archief om terug te kunnen zoeken wat er allemaal besproken is en als tool om ze er aan te herinneren dat er op een bepaald tijdstip iets moet gebeuren. Dit toegenomen gebruik van de functionaliteit van e-mail zorgt ervoor dat overloadgevoelens eerder plaatsvinden en dat er dus goed gekeken moet worden naar het ontwerp van e-mailsystemen.

Een artikel uit de New York Times uit 1997 onderschrijft deze conclusie. Te veel verstuurde of ontvangen e-mail zorgt er voor dat het risico gelopen wordt dan men de belangrijke berichten mist. Het systeem wordt langzaam door de ophoping van berichten en waardevolle schijfruimte gaat verloren. Personeel krijgt last van overloadgevoelens en het is aan het systeemontwerp om dit aan te pakken¹⁰.

In het artikel van Stevens uit 2000 blijkt dat de definitie van overload in het nieuwe millennium niet veel veranderd is. Het gaat alsnog om de hoeveelheid aanbod die de te verwerken hoeveelheid informatie overstijgt. Wat in dit artikel ook duidelijk naar voren komt is dat de oorzaak van overload en de daarbij behorende stressgevoelens voornamelijk ligt in het gebruik van e-mail. Mensen gebruiken bijvoorbeeld te makkelijk 'cc' en 'bcc' velden in mailberichten en berichten komen dus bij te veel ontvangers aan. Vaak zijn er e-mailprotocollen binnen organisaties, maar de controle of mensen zich ook aan deze protocollen houden ontbreekt over het algemeen. De conclusie van Stevens is dan ook dat er meer strategisch nagedacht moet worden over het gebruik van e-mail binnen organisaties. Dit vooral met betrekking tot protocollen en procedures die gevolgd dienen te worden bij het gebruik van e-mail¹¹.

Jeanne Allert geeft in haar artikel aan dat er ook andere mogelijkheden zijn om het gevoel van overload te verminderen. Zij geeft een 12-stappenplan aan dat gebruikers van e-mail aan het denken moet zetten over hun e-mailgedrag. Dit onderschrijft de conclusie die eerder getrokken is over het feit dat de manier waarom e-mail gebruikt wordt een van de grootste veroorzakers van het gevoel van overload is¹².

Ralph Adam signaleert dat werknemers met een erg efficiënte manier van e-mailafhandeling behoefte hebben aan een all-in-one tool voor het schrijven en verwerken van mailtjes, terwijl de meer traditionele gebruikers dit nog het liefst met

behulp van de gangbare e-mailcliënt doen. Dit opent wel voor het eerst de deuren voor een spraakherkenningssysteem om aan deze behoefte van efficiënte mailgebruikers te voldoen¹³.

Jenni Ingham heft in 2003 een uitgebreid onderzoek uitgevoerd in het Verenigd Koninkrijk over het gebruik van e-mail en de relatie tot information overload¹⁴. Uit haar onderzoek bleek dat het overgrote deel van communicatie binnen bedrijven tegenwoordig plaatsvindt via de mail. Waar vroeger een half uur nodig was om een brief op te stellen heeft men tegenwoordig gemiddeld 5 minuten nodig om een mailtje met dezelfde strekking te typen. Bijkomend nadeel is wel dat er veel meer spelfouten gemaakt worden in deze berichten (gemiddeld 57 spelfouten per 100 berichten¹⁵, maar dat men deze fouten accepteert en niet als zeer storend ervaart. In 1998 bleek al uit een onderzoek uitgevoerd door KPMG dat 27% van de ontvangers van bedrijfsmail vonden dat er ondraaglijk veel informatie aangeboden werd en nog eens 41% van de respondenten verwachtte binnen 5 jaar een ondraaglijke hoeveelheid informatie te ontvangen. Ingham concludeerde in haar onderzoek dat er veel winst gehaald kan worden uit het opstellen van een vaste procedure voor het omgaan met communicatie via e-mail. Zij legt de oorzaak van het overload gevoel dus niet bij de technologie, maar bij het gebruik van deze technologie.

In 2003 is wederom een enquête uitgevoerd over het gebruik van e-mail door David Dawley¹⁶. Uit de resultaten bleek dat het onzorgvuldige gebruik van e-mail grotendeels de veroorzaker was voor het overloadgevoel van managers. Saillant detail is dat de gebruikers van e-mail over het algemeen vonden dat zij zelf geen training nodig hadden voor het gebruik van mail, maar dat vooral hun collega's behoefte hadden aan training. Een typisch geval van 'de pot verwijt de ketel dat hij zwart ziet'. Wel geeft dit aan dat er meer tijd en moeite gestopt moet worden in het opzetten van vaste mailprocedures. Ook was een belangrijke conclusie dat het gevoel van overload afnam naarmate de gebruikers hoger opgeleid waren of erg ervaren waren in het gebruik van e-mail. Volgens Dowley zou er door bedrijven dus meer aandacht geschonken moeten worden aan trainingssystemen voor het gebruik van e-mail en dat er verder onderzoek gedaan moet worden of deze trainingssystemen het overloadgevoel ook kunnen laten afnemen.

McKeen et al. signaleerden in hun artikel uit 2004 dat de 'cc' en 'reply all' functie binnen e-mail een grote veroorzaker van overload bij midden en senior management is¹⁷. Deze overvloed aan binnenkomende e-mail zorgt ervoor dat een hoop managers overwerken en soms zelfs thuis werken om de elektronische communicatie af te handelen. Ook hun conclusie was dat het opstellen van e-mailprotocollen en het gebruik van meer geavanceerde technologieën binnen bedrijven de binnenkomende mail kon verminderen en zo overloadgevoelens bij managers grotendeel weg kunnen nemen.

In 2005 verschijnt de publicatie van Schuff et al. waarin het clusteren van berichten beschreven wordt als oplossing voor het verminderen van overloadgevoelens¹⁸. Ze vonden een sterk bewijs dat het clusteren van ontvangen berichten ervoor zorgt dat het gemakkelijker wordt voor gebruikers om berichten die van belang zijn te lokaliseren in hun inbox. Het systeem dat de auteurs voorstellen werkt volledig geautomatiseerd. Binnenkomende berichten worden gescand en worden gelijk in het cluster geplaatst dat het meeste overeenkomt vertoont met het bericht. De afhandeling van e-mail wordt voor de gebruiker simpeler waardoor de overload afneemt.

In hetzelfde jaar verschijnt ook een artikel van Betlotti et al. waarin ook gekeken wordt naar het automatisch filteren van binnenkomende e-mail op relevantie, belangrijkheid van afzender steekwoorden uit de inhoud¹⁹. Zij hebben de applicatie TaskMaster ontwikkeld die het filteren automatiseert. In dit artikel komt echter naar voren dat veel gebruikers de automatische filtersystemen niet vertrouwen en toch graag zelf de berichten zien, zodat ze zeker weten dat er geen belangrijke berichten als onbelangrijk bestempeld worden.

In 2006 verschijnt in het journal of Business and Technical Communication een artikel van Gil Thomas et al.⁵ waarin ook gekeken wordt naar de sociale impact die e-mail overload kan hebben op werknemers. Na drie verschillende studies naar het gebruik van e-mail blijkt dat drie hoofdkarakteristieken de hoofdoorzaak vormen voor overloadgevoelens en deze gevoelens ook beïnvloeden, te weten instabiele verzoeken, de druk om snel te antwoorden en de snelle verschuivingen van deelnemers binnen een discussie die via e-mail gevoerd wordt. In de periode van 1995 tot 2001 blijkt het e-mailverkeer al met 600% te zijn toegenomen en deze groei lijkt niet te stoppen. Als de capaciteit van ontvangers om informatie te verwerken toeneemt, zou het gevoel van overload moeten afnemen. Bij de auteurs leefde het gevoel dat inefficiënt en ineffectieve opbouw van de berichten een grote oorzaak zou zijn voor overloadgevoelens. Na analyse van de data bleek echter dat maar 3% van de 1727 onderzochte berichten opnieuw gelezen hoefde te worden om de strekking te begrijpen en dat 70% van de ontvangen berichten binnen 1 minuut gelezen konden worden. Wel bleek dat relatief korte mailtjes met een verzoek om informatie of een actie erg veel vervolgwerk konden veroorzaken. Dit droeg wel flink bij aan de overloadgevoelens bij de ontvangers.

Ook in 2006 verschijnt een uitgebreid artikel in het journal Communications of the Association for Information Systems van Ashish Gupta et al. met de titel "Mail Management - A Techno-Managerial Research Perspective"²⁰. In deze publicatie wordt onderzocht welke invloed e-mail heeft op het gedrag van ontvangers. Een opvallende conclusie is dat het gebruik van e-mail werknemers kan zorgen voor irrationeel gedrag en gedrag vergelijkbaar met ziektes als ADHD. De drang om elk moment van de dag te kijken of er nieuwe e-mail is en de wil om deze mailtjes ook direct te beantwoorden zorgt ervoor dat de productiviteit van werknemers afneemt. Ook signaleren zijn een achteruitgang van de taal door het toenemende e-mailverkeer. Afkortingen en acroniemen (bijvoorbeeld c u l8er) nestelen zich in het gewone taalgebruik zodat er sneller gereageerd kan worden op binnenkomende mail.

Het gebruik van laptops in vergaderingen kan ingedeeld worden in twee categorieën, namelijk 'meeting aligned use', bijvoorbeeld aantekeningen maken en 'non-aligned use', bijvoorbeeld het spelen van spelletjes of het sturen van e-mails. Ook dit heeft invloed op de prestaties van de werknemers. Gupta et al. concluderen dat werknemers zich aan moeten leren om maximaal twee maal per dag hun e-mail te checken om ervoor te zorgen dat de productiviteit gewaarborgd blijft en er niet teveel afleiding ontstaat door de continue stroom van binnenkomende mailtjes. Ook hopen ze dat er e-mail programma's zullen ontstaan die de gebruiker leren om met de mail om te gaan en het omgaan met de grote hoeveelheid binnenkomende berichten kunnen vergemakkelijken.

In 2007 is een onderzoek uitgevoerd door Wouter Spoelstra over het probleem van e-mail overload binnen de publieke sector in Nederland²¹. Hij heeft een casestudie uitgevoerd in de publieke sector. Hierbij werd inderdaad gedetecteerd dat het veelvuldige gebruik van e-mail een negatieve invloed had op de productiviteit van de managers. Over het algemeen werd hiermee omgegaan door e-mail met hoge prioriteit zo snel mogelijk te behandelen terwijl berichten met een lagere prioriteit behandeld werden zodra er tijd over was. Over het gebruik van VRS werd wisselend gereageerd door de ondervraagde managers. Er werd toekomst gezien in de technologie als er goed omgegaan kon worden met 'eehs' en 'aahs' in gesproken tekst, maar ook werd er sceptisch gereageerd met de opmerking dat er een té groot verschil zit tussen gesproken en geschreven taal. De auteur sluit af met de opmerking dat het gebruik van voice input in een bedrijfsomgeving een onderwerp is waar nog verder onderzoek aan besteed dient te worden.

In 2007 heeft Ruben Span een onderzoek over e-mail overload uitgevoerd bij een bedrijf uit de commerciële sector²². Ook hier waren de conclusies dat de hoeveelheid binnenkomende mail voor problemen bij managers kan zorgen. De oplossing voor dit probleem was volgens de geïnterviewden voornamelijk het geven van training met

betrekking tot het gebruik van e-mail binnen het bedrijf. Zoals één van de geïnterviewde managers zegt: "we krijgen training voor het gebruik van alles, behalve voor hetgeen dat we het meeste doen: e-mailen."

4.4 Tijdljn Spraakherkenning

In 1993 verschijnt het boek "Fundamentals of Speech Recognition", geschreven door Lawrence Rabiner en Bing-Hwang Juang waarin voor het eerst een theoretisch goed onderbouwde, technisch accurate en complete beschrijving van de basiskennis en ideeën staat die een modern systeem voor spraakherkenning door machines weergeeft.

In 1996 verschijnt het boek 'Using Speech Recognition' van J. Markowitz²³ waarin voor het eerst het gebruik van spraakherkenning wordt vergeleken met de traditionele manier van computer input. Zoals verwacht bleek uit de conclusies dat nieuwe gebruikers meer fouten maakten met het gebruik van de nieuwe technologie dan bij input via het keyboard. De gebruikers vonden het gebruik van spraakherkenning echter minder prettig, juist omdat er zoveel fouten gemaakt werden die achteraf weer verbeterd dienden te worden. Deze perceptie van de technologie veranderde niet als de typevaardigheid van de proefpersonen werd meegenomen in de vergelijking.

In 1998 wordt er een artikel gepubliceerd door Farzad Ehsani en Eva Knodt waarin commerciële systemen voor spraakherkenning beschreven worden²⁴. Ze behandelen vooral ViaVoice van IBM en NaturallySpeaking van Dragon. Deze systemen halen in eerste instantie een nauwkeurigheid van 60 tot 80%, maar deze getallen kunnen verbeterd worden door het systeem te trainen met het gebruik van de eigen stem. Hierdoor kan de foutenmarge verkleind worden tot 5%, maar iets meer conservatieve schattingen komen tot een foutenpercentage van 11 tot 13%. Andere factoren die van invloed zijn op de nauwkeurigheid van het systeem zijn de aanwezigheid van omgevingsgeluid, kwaliteit van de hardware (microfoons e.d.) die gebruikt worden. Een eis die gesteld wordt aan dit onderzoek is dat de spraakherkenningssystemen gebruikt kunnen worden op betaalbare PC platforms die voldoen aan de basis geheugeneisen.

D. Coniam concludeert in zijn onderzoek uit 1999²⁵ dat er wel rekening gehouden moet worden met de moedertaal van de gebruikers van spraakherkenningssystemen. Als de tekst gedicteerd wordt in de moedertaal maakt het gebruik veel minder fouten. Als een gebruiker eerst moet vertalen in een 'vreemde' taal (meestal Engels) dan loopt het herkenningpercentage zienderogen af.

Greg Alwang²⁶ beschrijft ook een nauwkeurigheid van herkenningssystemen in 1999 van 91 tot 95% als een gebruiker eerst de tijd neemt om het systeem te trainen. In voorgaande versies van de software varieerde dit percentage tussen de 80 en 90%. Een belangrijke conclusie van deze auteur was dat in 1999 het (blind) typen nog altijd sneller ging dan herkenning met gebruik van de software. Ook signaleerde hij een toename van het gebruik van spraakherkenningssystemen in bepaalde, gespecialiseerde bedrijfstakken als radiologie, politie, juridische en medische omgevingen.

In de publicatie van Tanya Goette uit 2000²⁷ blijkt dat er een aantal belangrijke voorwaarden zijn waar aan moet voldaan voordat een spraakherkenningssysteem succesvol geïmplementeerd kan worden in een bedrijfsomgeving. Zij geeft aan dat er geen onrealistische verwachtingen geschept moeten worden. Als er wonderen op het gebied van efficiëntie verwacht worden komen de gebruikers vaak bedrogen uit. Gebruikers denken alvorens de VRS te gebruiken vaak dat ze veel tijd aan het systeem moeten besteden voordat het voldoet aan hun verwachtingen. Als men direct resultaat verwacht en dit blijft uit dan haken veel gebruikers al af in de eerste fase. Goette maakt de vergelijking van het gebruik van Windows zonder muis. Het is zeker mogelijk, maar het zorgt voor veel frustratie tijdens het aanpassingsproces van de gebruiker. Werkgevers die hun werknemers in staat stellen om de tijd te nemen om de

herkenningssystemen goed te trainen zagen vaak een toename van tevreden gebruikers. Ook de training die nodig is om het systeem goed door te krijgen is erg belangrijk. Een gebrek hieraan zorgt gemiddeld voor een afhaakpercentage van rond de 40%. Een fout die ook vaak gemaakt werd is dat een werkgever, die succesverhalen te horen kreeg met betrekking tot de invoering van spraakherkenningssoftware bij bevriende bedrijven, automatisch uitging van het feit dat het in zijn organisatie ook nuttig zou kunnen zijn. In de praktijk blijkt dit echter niet waar te zijn. Dit voorbeeld geeft dan ook aan dat het heel belangrijk is om te kijken naar het organisatieontwerp en dat een bijpassend softwarepakket gekozen moet worden om het geheel te laten functioneren.

Uit het artikel van Haxer et al.²⁸ Blijkt dat er ook nadelen verbonden zijn aan het excessieve gebruik van spraakherkenningssoftware. Gebruikers starten met het gebruik van deze technologie omdat zij moeite hebben met het gebruik van keyboard en/of muis, of zelfs omdat ze de nieuwe technologie graag willen proberen. De meeste softwarepakketen die op de markt zijn vereisen een dicterende vorm van spreken. In dit artikel wordt voor het eerst gesproken over NaturallySpeaking van Dragon, dat gebruikt maakt van natuurlijke spraakpatronen en waardoor meer een gesprek met de computer gevoerd kan worden dan dat er via allerlei commando's interactie plaatsvindt. Een negatief bijeffect van deze spraakherkenning ontdekken ze dus echter ook. Ze komen er achter dat er erg veel gesproken dient te worden en dat dit bij sommige gebruikers tot stemproblemen kan leiden.

In 2001 verschijnt ook het artikel van Helmut Mangold²⁹, verbonden aan het research center van DaimlerChrysler. Hij onderzoekt het gebruik van spraakherkenning in de auto-industrie. De belangrijkste conclusie van zijn artikel is ook te gebruiken voor het gebruik van spraakherkenning in het algemeen. Hij ontdekt dat het gebruik van spraakherkenning zeker een toekomst heeft, maar dat het ontwerp van systemen om de achtergrondgeluiden en ruis te onderdrukken hierbij een hele belangrijke rol gaan spelen. Hij signaleert de vooruitgang die spraakherkenningstechnologie al gemaakt heeft en ziet, ondanks dat de systemen nog verre van optimaal werken, de potentie voor toekomstig gebruik. Alleen zal er zeker nog verder onderzoek gedaan moeten worden naar de robuustheid van de technologie, te weten het goed om kunnen gaan met veranderende omgevingsparameters als omgevingsgeluid en variatie in stemvolume en stemgeluid van de gebruikers.

De toekomst die in series als Star-Trek al werd voorgeschoteld wordt inmiddels steeds meer werkelijkheid. Charles Lowe³⁰ ziet de ontwikkeling van erg dure softwarepakketen die alleen draaien om zeer snelle computersystemen in de jaren 80 tot erg betaalbare spraakherkenningssoftware die draait op de standaard PC's voor huis-tuin-en-keukengebruik. Hij zegt echter wel dat er nog een verandering in denken bij de gebruikers moet plaatsvinden. Het gebruik van keyboard en muis is gemeengoed, maar dat gebruikers via spraak interactie hebben met een computer is voor veel mensen nog niet normaal. Het feit dat het voor mensen erg natuurlijk is om via spraak te communiceren zal de overstap naar spraakgestuurde machine-interactie vergemakkelijken. Charles Lowe signaleert in een onderzoek bij studenten aan de Florida State University dat studenten die het omgaan met spraakherkenningssoftware onder de knie hebben veel grote tekstverwerkingssnelheden halen dan collega's die het toetsenbord gebruiken voor de input. De weerstand die nog bestaat bij gebruikers zal snel weggenomen worden als kinderen opgroeien met deze technologie. Softwareontwikkelaars zullen deze techniek sowieso op de markt brengen en de markt zal zich zeer waarschijnlijk aanpassen aan deze nieuwe technologie. Hij vraagt zich ook af of in de toekomst onze taal zich ook zal aanpassen aan dit soort systemen, maar dat zal in de toekomst nog moeten blijken. Hij voorziet een toenamen van het gebruik van 'spreektaal' in communicatie hij geeft aan dat de jeugd de toekomst bepaald en het zogenaamde 'freespeaking' ook steeds meer de norm zal worden in onze moderne communicatie.

Harris en Biermann³¹ halen een onderzoek aan over of het gebruik van spraakherkenning om door menu's te navigeren sneller is dan het gebruik van een muis. Uit dit onderzoek blijkt dat het gebruik van spraakherkenning pas sneller wordt bij menu's die uit meer dan 64 stappen bestaan. Bij het navigeren door kleinere menu's wint de muis het van de spraakherkenning. Een bijkomende vondst was dat bij menu's bestaand uit korte woorden meer fouten door de software gemaakt werden dan bij het gebruik van lange woorden in de menustructuur.

Een onderzoek van Deng et al.³² bij Microsoft research, een project genaamd Dr. Who, concludeert dat het multimodaal gebruik van gesproken taal als interfacemethode de bruikbaarheid van computers aanzienlijk kan verhogen omdat spraak en GUI complementaire sterktes hebben. De reden waarom spraak niet als mainstream interfacemethode wordt gebruikt ligt in het feit dat in de taal veel dubbelzinnige woorden bestaan, de gebruiker veel over het gebruik van de technologie (specifieke commando's) moet onthouden en dat de technologie nog verre van foutloos werkt. De conclusie die uit dit onderzoek getrokken kan worden is dat in de toekomst een combinatie van verschillende input-technieken, zoals het gebruik van keyboard en muis, grafische displays en het gebruik van spraakherkenning en overige audio- en videotekniken, in de toekomst de interactie van gebruikers met technologie dramatisch zal vergemakkelijken.

In het artikel van Amit Mehta et al.³³ uit 2003 wordt geschat dat in 30% van de radiologiepraktijken gebruik maken van spraakherkenningstechnologie of van plan zijn om dit te gaan gebruiken. Dat de gebruikers de tijd namen om het systeem aan hun stem en woordgebruik te laten wennen werd als zeer belangrijk ervaren. Het produceren van radiologie-rapporten (het bespreken van X-Ray foto's) ging met gebruikmaking van spraakherkenning tot 10 maal sneller. Op de afdeling die werd onderzocht namen de kosten voor de rapportproductie ook met \$350.000 af. Deze besparing werd zelfs al met een gedeeltelijke implementatie van de technologie gerealiseerd. Een belangrijke conclusie was echter wel dat de lasten van de rapportproductie nu wel bij de radiologen kwam te liggen. Eerder werden in een voicerecorder ingesproken beoordelingen door typistes uitgewerkt, maar bij het gebruik van spraakherkenning moesten de correcties (gemiddeld 10 tot 15 belangrijke correcties per rapport) door de radiologen zelf worden uitgevoerd. Een belangrijke conclusie is dat de productiviteit van de radiologen bij het gebruik van spraakherkenning dus afnam. Het belang van spraakherkenning voor de toekomst werd echter wel ingezien en de schatting van de auteur was dat binnen 5 tot 10 jaar na het onderzoek spraakherkenning in bijna alle radiologie-afdelingen gemeengoed zou zijn.

Larson en Mowett³⁴ hebben een onderzoek gedaan naar het gebruik van spraakherkenningssoftware en vooral naar hoe mensen fouten die door de software worden gemaakt corrigeren. De meeste gebruikers proberen eerst met behulp van de software de fouten te corrigeren, maar aangezien dit vaak niet werkt gaan de meeste gebruikers na deze eerste poging over op het handmatig corrigeren van de tekst. Voor het aantal keer dat een gebruiker probeert de fout via de software te corrigeren door opnieuw in te spreken wordt door de auteurs de term 'spiral depth' geïntroduceerd. In de helft van de tijd dat de software gebruikt wordt gaan gebruikers tot een 'spiral depth' van 3 en 25% van de tijd tot een 'spiral depth' van 4. Efficiënte gebruikers bleken tot een maximale 'spiral depth' van 2 te gaan en daarna over te gaan op een andere correctiemethode. Bij het navigeren tussen de fouten met behulp van de software werden nog meer fouten gemaakt waardoor er een waterval aan fouten ontstond. Uiteindelijk bleek het gebruik van de software voor foutverbetering 10x zo lang te duren dan het opnieuw intikken van de tekst via het keyboard. De conclusie van dit onderzoek was dan ook dat de barrière om spraakherkenningssoftware te gebruiken in plaats van traditionele tekst-input vooral ontstaat door omslachtige methode om de gemaakte fouten te verbeteren. Hier zou bij de softwareontwikkelaars dan ook meer de nadruk op

moeten liggen, waar deze nadruk nu vooral ligt in het verfijnen van de herkenning en dus het terugdringen van de foutenmarge.

Uit het onderzoek van Michael Grasso uit 2003 onder 31 auteurs van papers over het gebruik van spraakherkenning in medische applicaties³⁵ blijkt dat de trends in de spraakherkenningssoftware van de laatste jaren bestaan uit onafhankelijkheid van gebruiker en woordenboeken (waardoor het aanleren van de specifieke gebruikersstem aan de software overbodig wordt) en continue spraakherkenning. In het medische veld worden deze innovaties voornamelijk toegepast bij het gebruik van dicteersystemen voor het produceren van rapporten bij radiologie, pathologie en endoscopie. Een enquête uitgevoerd door Grasso geeft aan dat men erg optimistisch is over het gebruik van spraakherkenningssoftware binnen het medische vakgebied. Elk jaar neemt de voorkeur voor deze technologie nog toe en bijna alle deelnemers aan de enquête verwachten dat hun organisatie spraakherkenningstechnologie in de (nabije) toekomst meer zullen gaan gebruiken.

Uit een onderzoek van James Hartley uit 2003³⁶ bleek dat er geen significante verschillen bestonden tussen de gemiddelde lengte van geschreven teksten, hoeveelheid paragrafen en hoeveelheid gebruikte zinnen tussen via VRS gegenereerde tekst en getypte teksten. Ook zaten er niet significant meer spellingsfouten in gedikteerde teksten. De kwaliteit van de tekst bleef dus in stand. Wel viel op dat er bij het dicteren van tekst kortere zinnen gebruikt worden en dat de ik-vorm vaker gebruikt wordt. Uiteindelijk concludeerde Hartley dat het gebruik van spraakherkenning een marginaal effect had op de kwaliteit van de 'opgeleverde' tekst, maar dat het een grote impact had op de beleving van de schrijver. Het gebruik van de software werd initieel als lastig ervaren en men vond dat er een hoop oefening nodig was voordat alles naar behoren werkte. En zelfs na veel oefenen vond men het moeilijk om goed lange en complexe zinnen te blijven gebruiken. Opvallend was wel dat de gebruikers van de software niet aangaven dat ze graag terug wilden naar het gebruik van het keyboard. Men vond het fijn dat door het gebruik van spraakherkenning in plaats van typen de handen gebruikt konden worden om tegelijkertijd een andere taak uit te voeren. Wat verder nog benadrukt dient te worden is dat geen van de gebruikers van de software het keyboard helemaal niet meer gebruikte. Duidelijk werd dat er een hybride vorm van tekstinput ontstond waarbij er zowel van de software als van het keyboard gebruik werd gemaakt om tekst in te voeren.

Volgens een onderzoek uit 2003 van Rieger³⁷ blijkt dat spraakherkenning een economisch verantwoorde optie kan zijn voor bedrijven die secretariële uitgaven willen terugdringen. Ook ziet hij spraakherkenning als optie die het gemak verhoogt van mensen die liever dictatie gebruiken in plaats van typen met een keyboard.

Uit het onderzoek van Robert Issenman³⁸ blijkt dat de tijd die artsen nodig hebben om een rapport op te stellen (schrijven en corrigeren) met het gebruik van VRS 200% langer is dan dat een ervaren transcriptionist de tekst uittipt en vervolgens importeert in het medisch dossier (9 minuten voor de arts in plaats van 3 minuten voor de transcriptionist). Als de licentiekost voor de software en het hogere uurloon van de arts ten opzichte van de transcriptionist meegenomen worden in de vergelijking blijken de totale kosten van het gebruik van VRS 100% hoger te liggen. Ook signaleert hij dat een hoop artsen geen zin en tijd hebben om het systeem te trainen. Uit eerdere literatuur blijkt dat VRS al succesvol wordt toegepast binnen radiologie-afdelingen en dat de tevredenheid significant toeneemt naarmate er gespecialiseerde woordenboeken door de VRS worden gebruikt. Dit lijkt te verklaren doordat er veel gestandaardiseerde termen binnen dit medische veld worden gebruikt in de rapportages. Commerciële softwareproducenten propageren dan ook een hybride gebruik van de software, waar transcriptionisten de conceptrapporten van de artsen corrigeren. Dit zou kostbare tijd van de artsen kunnen besparen, maar dan moet er wel uitgegaan worden van het feit

dan een transcriptionist begrijpt wat er bedoeld wordt met een fout in het conceptrapport.

Bryan Bergeron³⁹, verbonden aan de afdeling Health Science and Technology op het Massachusetts Institute of Technology (MIT) heeft honderden artikelen geschreven over het gebruik van toegepaste technologieën binnen de zorg. In zijn artikel uit 2004 blijkt dat het gebruik van VRS in de zorg toeneemt, maar niet voor elke praktijk handig kan zijn. Kleine praktijken kunnen profiteren, maar voor grote praktijken kan er beter gebruik gemaakt worden van (goedkopere) transcriptionisten. Hij signaleert namelijk dat een transcriptionist sneller teksten kan corrigeren terwijl hij luistert naar de tekst die ingesproken is op een dictafoon dan dat een arts ze kan produceren en corrigeren met VRS. Er wordt ook geld bespaard omdat de transcriptionist niet het gehele rapport hoeft uit te typen, maar alleen de correcties hoeft uit te voeren naar aanleiding van de opgenomen tekst. Ook hier komt dus naar voren dat een hybride vorm het meeste voordeel kan opleveren. Hij haalt ook nog aan dat het toenemende gebruik van offshore transcription services in India en Pakistan in de toekomst de kosten voor het corrigeren en uittypen nog verder kunnen dalen.

Uit het onderzoek van Atif Zafar et al.⁴⁰ uit 2004 blijkt dat zowel bij het gebruik van VRS als bij het uittypen van gedicteerde tekst door transcriptionisten nog veel fouten gemaakt worden. Wat opviel in dit onderzoek was dat bij het gebruik van VRS de fouten vaak gecorrigeerd konden worden door de context goed door te lezen. De auteur schrok wel van de hoeveelheid fouten die bij beide systemen ongecorrigeerd in de rapporten bleven staan. Een klein aantal van deze fouten zou ook invloed gehad hebben op de soort zorg die de patiënt gekregen zou hebben. Een belangrijke aanbeveling van de auteurs was dan ook om VRS systemen uitgebreid te testen voordat ze toegepast worden in een echte klinische omgeving.

Binnen de medische industrie schatten de CIO's en de artsen zelf dat 50% van alle artsen binnen 5 jaar een vorm van VRS zullen gebruiken. De grootste problemen die Terheyden⁴¹ signaleert hangen samen met spraakverschillen tussen gebruikers, het oplossen van dubbelzinnigheid van sommige woorden, hoe om te gaan met achtergrondgeluiden, en interpunctie- en grammaticaregels die vereist zijn voor definitieve rapportages. Hij splitst de kosten die gemaakt worden voor het maken van documentatie op in de kosten van de tijd die de arts er aan besteedt, transcriptiekosten, en de kosten die samenhangen met de tijd dat de documenten niet beschikbaar zijn voor het leveren van zorg aan patiënten. Ook geeft hij aan dat de kosten die gemoeid zijn met het gebruik van externen, zoals transcriptionisten, editors, en administratieve kosten die samenhangen met archiveren, printen en analyseren van rapporten meegenomen moeten worden in het totale kostenplaatje. Met een toenemend gebruik van VRS nemen de kosten samenhangend met de tijd die de arts in het rapport neemt toe, maar nemen de overige kosten af. Het is aan de instelling om te bepalen waar het break-even punt ligt en of het rendabel is om VRS te gebruiken. In de toekomst zullen door toename van de accuratesse van de VRS systemen de kosten voor rapportproductie waarschijnlijk nog verder dalen.

In een studie van Feng et al.⁴² uit 2005 blijkt dat training de efficiëntie van gebruikers van VRS systemen flink doet toenemen. Waar een beginner voor een bepaalde taak gemiddeld 41 minuten nodig had werd deze tijd teruggebracht tot 25 minuten bij een soortgelijke taak na 7x oefenen. Wat wel opviel is dat van deze 25 minuten gemiddeld 13 tot 16 minuten werd besteed aan het navigeren en het verbeteren van fouten. Hier is dus nog heel veel winst te halen bij de ontwikkeling van nieuwe VRS pakketten.

In 2005 hebben Robert DeRosier en Ruth S. Farber⁴³ een onderzoek uitgevoerd over het gebruik van VRS door mensen met een handicap. Ook hier werd duidelijk dat de testpersonen tevreden waren over de software, maar dat er nog veel voor verbetering vatbaar was. In het geval van mensen met een handicap, die niet via de standaard

inputmethodes gebruik kunnen maken van de computer, opent de VRS deuren die eerder gesloten bleven. Ook uit deze studie bleek dat er nog verder onderzoek nodig was naar de correlatie van de gebruikerstevredenheid en de hoeveelheid training voor het gebruik van de VRS systemen.

Keith White⁴⁴ detecteert een revolutie in de manier waarop radiologierapporten gemaakt worden. Deze revolutie wordt gedreven door twee factoren, namelijk het verminderen van de afdelingskosten en het verlagen van de doorlooptijden van rapporten. Op financieel gebied is er een vrij grote investering nodig om met VRS aan de slag te gaan (aankoop van de software en de desktop computers waar het gebruikt gaat worden, maar zeker ook de servers waar de software op draait zodat het op meerdere locaties te gebruiken is binnen de afdeling), en bij het geven van training en het onderhoud van de software zijn ook nog flinke kosten gemoeid. Radiologen die de software gaan gebruiken zijn in eerste instantie ook niet onder de indruk van de technologie, omdat ze bij aanvang vaak meer tijd nodig hebben om de rapporten te genereren. Ook zitten er als de VRS voor het eerst gebruikt gaat worden nog fouten in de rapporten en moet er dus vrij veel tijd in het corrigeren van deze rapporten worden gestoken. Tevens heeft de invoering van VRS een zijdelingse negatieve werking. De interactie tussen de radiologen onderling en tussen de verschillende schakels op de afdeling in het bijzonder neemt af aangezien er meer achter het scherm geregeld wordt in plaats van face-to-face. Het feit dat de afdeling ook erg afhankelijk is van de technologie en elektronica en dus bij een stroomstoring of systeemcrash in grote problemen kan komen moet meegenomen worden in de beslissing om VRS te gaan gebruiken. Keith White ziet echter zeker grote voordelen aan de implementatie van VRS binnen de afdeling radiologie. De hogere kosten voor implementatie worden zijns inziens volledig tenietgedaan door de voordelen. De operationele afdelingskosten dalen significant bij gebruik van VRS en het verlies aan productiviteit van de radiologen wordt makkelijk gecompenseerd. Een groot tweede voordeel is dat de productietijd van de rapporten dramatisch afneemt. De tijd van dicteren tot handtekening liep van 42 uur voor implementatie van VRS af tot 5 minuten na implementatie. Binnen de zorg kan deze tijdwinst ook mensenlevens redden. De rapportages van radiologen zijn bij verder diagnosticeren en het bepalen van een zorgtraject sneller beschikbaar en hebben dus ook meer invloed op het beslissingsproces van de artsen. Hierdoor voelen radiologen zich meer gewaardeerd en dit weegt vaak op tegen het gevoel van verlies van productiviteit dat ze bij het begin van de implementatie ervaren.

Samengevat zegt White dat de technologie van spraakherkenning inmiddels volwassen geworden is en het een zeer nuttige bijdrage kan leveren aan het verlagen van operationele kosten en verbeteren van de zorg. Training is erg belangrijk zodat er geen fouten in de rapportages verschijnen die potentieel levensbedreigend kunnen zijn, maar dat de radiologie voordelen kan halen bij het gebruik van VRS is voor hem evident.

Stephen Manes⁴⁵, een bekende columnist voor technologiebladen in de Verenigde Staten, schrijft in een column in 2005 in PC World dat Dragon NaturallySpeaking een VRS pakket is dat de laatste jaren flink wat verbeteringen doorvoert die de software steeds beter en gebruiksvriendelijker maakt. Het corrigeren van fouten wordt steeds makkelijker en kost vooral steeds minder tijd. Hierdoor neemt het gebruikersgemak van de software toe en hij signaleert dat dit softwarepakket moet worden beschouwd als een hulpmiddel om werk achter een computer op een andere manier sneller en beter te doen.

In een publicatie uit 2005 van de radiologen Richard L. Morin en Steve G. Langer⁴⁶ blijkt dat ook deze auteurs een grote toekomst zien voor spraakherkenningsystemen binnen de radiologie. Het feit dat het in 2005 al moeilijk was om überhaupt medische transcriptionisten te vinden en vervolgens als werknemers te behouden doet hun vermoeden dat in de toekomst alleen nog maar meer gebruik zal worden gemaakt van VRS binnen de radiologie. De nadruk komt te liggen op de snelle productie van rapporten, waar nauwelijks fouten in mogen zitten. Op deze manier is namelijk de

veiligheid van de patiënten gewaarborgd en kunnen er interpretaties gedaan worden van rapporten en foto's die van hoge kwaliteit blijven. Een opvallende waarneming is dat het de auteurs opviel dat een hoop gebruikers hun manier van spreken aanpassen aan de software. Zij geven terecht aan dat dit eigenlijk andersom zou moeten zijn.

Sixto Ortiz Jr.⁴⁷ heeft in een editorial voor Tech&Trends magazine een artikel geschreven over de commerciële toepasbaarheid van spraakherkenning in een bedrijfsomgeving. Hoewel volgens de auteur een intelligent gesprek met een computer nog niet mogelijk is komt het met de VRS van tegenwoordig wel dichtbij. Zijn conclusie is dat NaturallySpeaking van Dragon het beste pakket is dat momenteel op de markt beschikbaar is. Een aantal gebruikers (artsen en advocaten die door de auteur zijn geïnterviewd) geeft aan dat het softwarepakket erg goed werkt, maar wel erg veel training nodig heeft om volledig gebruik te kunnen maken van de mogelijkheden die het pakket biedt. Ook zal de manier van dicteren, vooral op het gebied van intonatie, spreeknelheid en uitspraak enigszins aangepast moeten worden om een optimaal resultaat te bereiken. Maar de auteur concludeert dat VRS voor de toekomst erg veel potentie heeft.

Julie Park et al.⁴⁸ hebben voor de universiteit van Auckland, Nieuw Zeeland, een kort onderzoeksproject uitgevoerd om te kijken of het mogelijk was om met behulp van VRS transcripties te maken. Uit het onderzoek bleek dat de training van gebruikers van de software twee kanten had. Ten eerste moesten de gebruikers leren omgaan met de software, wat voor mensen zonder IT achtergrond soms al moeilijk genoeg was. Vervolgens moest het programma ook wennen aan de gebruiker. Door fouten die gemaakt werden te verbeteren went de software ook aan het taalgebruik van de gebruiker waardoor de foutenmarge zienderogen afneemt. Om een voldoende nauwkeurigheid te bereiken was gemiddeld 10 tot 20 uur gebruik en correctie van tekst nodig. De meeste gebruikers waren echter al tevreden over het bereikte resultaat na 12 uur gebruik. Getrainde VRS zorgde uiteindelijk voor een besparing van ongeveer 50 euro per uur tekst dat uitgetypt moest worden t.o.v. gebruik van een secretaresse bij gelijkblijvende transcriptiekwaliteit. En gesignaleerd nadeel van het gebruik van VRS is dat gebruikers toch vaak gefrustreerd raken als de software niet naar behoren werkt en gebruikers kregen zelfs last van hun stem als ze de hele dag gewerkt hadden met VRS. Ook blijft de software vaak zelfde fouten maken die moeilijk te verbeteren zijn. Dit hoeft geen probleem te zijn tenzij de verkeerde herkende woorden erg belangrijk zijn voor de context.

Concluderend bleek uit het onderzoek dat het gebruik van VRS erg nuttig kan zijn in taken waar veel gesproken tekst uitgewerkt moet worden. De kwaliteit van het eindproduct komt overeen met de kwaliteit die bereikt wordt bij het gebruik van een ervaren transcriptionist. En hoewel er zeker nadelen waren vonden vooral proefpersonen met manuele handicaps of een trage typsnelheid de software erg bruikbaar. Zelfs na dit onderzoek zijn een aantal gebruikers meer gebruik gaan maken van VRS dan voordat zij aan dit onderzoek mee hadden gewerkt.

Ook binnen het defensie-apparaat wordt onderzoek gedaan naar het gebruik van spraakherkenning. Fang Chen⁴⁹ heeft hier een stuk over gepubliceerd in 2006. De argumenten die in het voordeel van het gebruik van spraakherkenning spreken zijn:

1. Spraak is de natuurlijk manier van informatie overdracht
2. VRS zorgt ervoor dat gebruikers hun handen en ogen vrij hebben en maximaal kunnen benutten.
3. De capaciteit om informatie te verwerken neemt toe
4. Er wordt een scheiding aangebracht in de informatie en de informatiebronnen
5. De kracht en nauwkeurigheid van VRS neemt erg snel toe
6. Vermindering van de belasting op personeel voor het uitvoeren van simpele gestructureerde taken

De communicatie binnen het leger leent zich uitstekend voor het gebruik van VRS. Het leger maakt gebruik van korte, gestandaardiseerde commando's, gestructureerde taken

en goed getraind personeel. Er zijn echter ook twee grote nadelen, te weten de onbetrouwbaarheid van de software en de lage fouttolerantie. Het feit dat er een fout gemaakt kan worden die grote consequenties kan hebben zorgt voor extra stress bij de gebruikers, wat vervolgens weer een negatieve invloed kan hebben op de herkenningsnauwkeurigheid. De nauwkeurigheid van 95% en beter die nu al gehaald kan worden geeft aan dat er zeker een toekomst is weggelegd voor het gebruik van VRS bij defensie, maar de verwachting van de auteur is dat het vooral gebruik zal worden voor niet-kritische onderdelen waar fouten geen grote gevolgen kunnen hebben.

Uit een onderzoek van Giovanni Cannata⁵⁰ uit 2006 blijkt dat tekst op papier bij een hoop gebruikers nog een niet te vervangen rol speelt binnen de informatieverwerking. Tekst op een scherm voldoet voor een hoop gebruikers nog niet zo goed als uitgeprinte tekst, en de vraag is dus of dit invloed kan hebben op het toekomstige gebruik van spraakherkenning, waarbij tekst dus voornamelijk digitaal aangeleverd en verwerkt wordt in plaats van op papier.

Linda Dailey Paulson⁵¹, tech-schrijfster in het Amerikaanse blad "Computer" schrijft in 2006 een stuk over de toekomst van VRS. Ze heeft interviews met een aantal experts op het gebied van VRS en met een aantal chipproducenten. Het overheersende gevoel is dat spraakherkenning pas echt gebruikt zal worden als het onderdeel gaat uitmaken van de hardware. Vroeger waren er aparte kaarten nodig voor videobewerking en -decoding en tegenwoordig zit deze functionaliteit al ingebakken in de hardware. VRS leunt zwaar op de processorcapaciteit van de PC en dit is een van de oorzaken dat de vraag naar VRS nog niet groot is. Als het functionaliteit wordt van de hardware en dus standaard onderdeel uitmaakt van een computer zal spraakherkenning pas veel gebruikt gaan worden.

Mike Wald en Keith Bain⁵² hebben in 2007 een onderzoek uitgevoerd onder 44 gehandicapte studenten die moeite hadden met het maken van aantekeningen door hun handicap. Het doel was een document via VRS te produceren dat identiek zou zijn aan het document dat gemaakt werd met behulp van input van een keyboard. Ze hebben gekeken of VRS een optie zou zijn om de kwaliteit van leven voor deze studenten te verbeteren. Deze vraag kon positief worden beantwoord. Ook kwamen zij met een aantal eigenschappen waaraan een succesvol spraakherkenningssysteem in een ideale situatie zou moeten voldoen.

- Volledig transparant zijn voor de 'spreker' (er is geen training nodig en de spreker hoeft geen microfoon te dragen)
- Volledig transparant zijn voor de 'luisteraar' (de gebruiker hoeft geen speciale apparatuur mee te dragen)
- Van een willekeurige spreker de stem herkennen (ongeacht accent, verkoudheid e.d.)
- Elk woord in elke context herkennen (dus ook of het een commando is of een beschrijving)
- Herkennen en overbrengen van houding, interesse, emotie en toon
- De spreker herkennen en dus herkenning wie het is en waar hij/zij zich bevindt
- Om kunnen gaan met elk type en niveau van omgevingsgeluid, spraakqualiteit en -volume

John Binder⁵³, een columnist in het blad Systems en Software heeft voor een column in 2007 gebruik gemaakt van de VRS van Nuance, te weten NaturallySpeaking versie 9.5. In zijn column is hij blij verrast over de kwaliteit van de software. De "Nothing but Speech" technologie van Nuance viel hem vooral op. Dit zorgt ervoor dat de 'uhm' en 'euh' termen die vaak in een gesproken verhaal geuit worden uit de tekst gefilterd worden. De column, geproduceerd met de software, kwam snel tot stand, en in elke paragraaf hoefden gemiddeld twee woorden veranderd te worden na teruglezen. De conclusie van de auteur was dat na voldoende training en oefening de software bijna foutloos zou moeten werken. Hij merkt terecht op dat het leren typen ook veel tijd en

moeite gekost heeft en dat dit voor gebruikers van VRS geen groot probleem zou moeten zijn. Ook merkt hij wel op dat de software van Nuance de muis overbodig kan maken, maar dat het gebruik van keyboard en muis in combinatie met de software het meest efficiënt werkt.

In 2007 schreef Vic Keegan⁵⁴, een van de belangrijkste tech-schrijvers van de Guardian in Engeland, een opiniestuk in zijn krant. Hij beschrijft hierin het gebruik van VRS en concludeert dat het gebruik van VRS vooral nut kan hebben bij het maken van persoonlijke aantekeningen. Hij heeft aan dat de technologie nog wel foutjes maakt, maar dat deze niet storend zijn in stukken tekst voor privé gebruik. Ook is hij benieuwd hoe de toekomst voor VRS er uit ziet, aangezien de technologie met de dag beter wordt.

In 2007 verschijnt in het journal Health Management Technology een artikel⁵⁵ over spraakherkenning en transcriptie. Hierin is Laura Cardio, de Health Information Supervisor van Dean Health Systems, een samenwerkingsverband van meer dan 60 zorgverleners uit verschillende disciplines, aan het woord. Binnen dit zorgcollectief, waar meer dan 500 artsen bij aangesloten zijn maken ongeveer 350 zorgverleners en 45 transcriptionisten intensief gebruik van VRS. Er wordt binnen dit collectief niet altijd gebruik gemaakt van aparte microfoon om met de software te werken. De verbetering in technologie heeft ervoor gezorgd dat zelfs via de telefoon gebruik gemaakt kan worden van het spraakherkenningssysteem. Doordat de VRS gebruik maakt van medisch taalgebruik modellen is de accuratesse flink toegenomen. Binnen Dean Health Systems worden nu maandelijks rond de 22.000 rapporten gegenereerd met behulp van VRS. Dit is ongeveer 1/3^e van het totaal aantal rapporten. Buiten het voordeel dat de artsen behalen blijkt dat de transcriptionisten ook een stuk productiever worden. Gemiddeld neemt de productiviteit van deze werknemers met 40% toe, met uitschieters naar 150%. Dit alles heeft tot nu toe een besparing van 300.000 dollar op het budget voor outsourcen geleid.

Douglas J. Quint MD⁵⁶ schrijft in een opiniestuk in het blad voor "The American College of Radiology" uit 2007 over spraakherkenningssystemen. Hij geeft aan dat er zeker nadelen kleven aan deze technologie, zoals afname van de productiviteit van radiologen, afnemende kwaliteit van rapporten, een negatief effect op het onderwijs aan artsen in opleiding en het verschuiven van de kosten van (goedkope) medisch transcriptionisten naar (dure) doktoren. Voordat deze nadelen worden weggenomen is het nog niet aan te raden om gebruik te maken van VRS binnen de radiologie, ook al zijn er zeker voordelen bij het gebruik van VRS bij het genereren van rapportages. Het belangrijkste voordeel is dat de rapportages sneller en makkelijker beschikbaar zijn, maar dit weegt volgende de auteurniet op tegen de nadelen.

In 2008 verschijnt in het International Journal of Speech Technology een artikel van Mike Wald et al.⁵⁷ over het gebruik van VRS door dove en slechthorende studenten. Colleges worden met behulp van VRS op schrift gezet zodat ook de studenten met deze handicap gebruik kunnen maken van deze vorm van onderwijs. De belangrijkste conclusie was dat deze systemen zeker kunnen bijdragen, maar dat er nog flink wat kanttekeningen te plaatsen zijn. Het editen van de teksten is erg omslachtig en neemt erg veel tijd in beslag. Ook de foutenmarge baart nog zorgen. Maar dat dit soort systemen in de toekomst zeker een positieve invloed kunnen hebben is zeker.

Ook in 2008 verschijnt in het International Journal of Speech Technology een artikel van András Kocsor en Gábor Gosztolya⁵⁸ over de mogelijkheden om het spraakherkenningproces te versnellen. Volgens de auteurs is niet alleen de accuratesse belangrijk, maar ook de vertaalsnelheid. Ze hebben een aantal technieken geëvalueerd en er bleek dat een combinatie van deze technieken het herkenningproces tot 12 keer sneller kon maken dan de gangbaar gebruikte technieken op zichzelf.

In 2008 valt op te maken uit een artikel van Mika P. Koivikko et al.⁵⁹ dat gepubliceerd is in het journal of Digital Imaging dat het invoeren van spraakherkenning bij het creëren van (medische) rapportages, de report turnaround times (RTT's) flink doet afnemen. Waar eerder alleen VRS beschikbaar was in de pan-Europese talen (Engels, Frans, Duits) zijn tegenwoordig ook pakketten beschikbaar in overige talen. Zelfs het Fins, wat voor VRS een uitdaging vormt door de punctuatie en uitspraak, vormt tegenwoordig geen probleem meer. In de ziekenhuizen in Europa wordt over het algemeen gebruik gemaakt van Philips SpeechMagic binnen de klinische applicaties die opereren met spraakherkenning. De RTT's van het Töölö ziekenhuis in Helsinki namen na ingebruikname van VRS af met 81%. Het aantal rapportages dat al binnen een uur beschikbaar was nam ook significant toe, van 26% tot 58%. De 'ouderwetse' manier van rapporteren, het inspreken op een cassettebandje en dat laten uittypen, werd nog steeds aangeboden naast het gebruik van VRS. Na invoering van VRS nam het aantal rapporten dat via het oude systeem werd geproduceerd af van 100% tot 12%. Het gebruik van VRS werd dus door het overgrote deel van de gebruikers omarmd.

De training voor het gebruik van VRS werd verzorgd door één radioloog van het ziekenhuis. De radiologen vonden het vrij simpel om te wennen aan het gebruik van de VRS en over het algemeen was 10 tot 15 minuten training per nieuwe gebruiker genoeg. Na de initiële training waren ervaren collega's bereid om vragen die naderhand opkwamen te beantwoorden. Concluderend namen niet alleen de doorlooptijden van de rapporten af, maar het gebruik van de online VRS rapportages deed de integrale patiëntenzorg verbeteren.

Kimberley Voll et al.⁶⁰ hebben in een paper uit 2008 uit het journal of Digital Imaging 2700 rapportages van het Canada Diagnostic Centre (CDC) gemaakt met behulp van de VRS Dragon NaturallySpeaking versie 7.3 geanalyseerd. Het resultaat was een post-speech-recognition foutendetectiesysteem. Statistische natuurlijke taal analyse kan het overgrote deel van vreemde lettercombinaties en foutief taalgebruik opsporen. Door deze inconsistenties te markeren kan de tijd die radiologen nodig hebben om hun met VRS geproduceerde rapporten te corrigeren significant doen afnemen.

De auteurs concluderen dat er een flinke potentie is voor VRS om radiologierapporten te produceren, maar dat de technologie momenteel nog een erg zwak en inconsistent alternatief is voor de traditionele menselijke transcriptiemethode. Dit komt vooral door de in de praktijk waarneembare magere nauwkeurigheid bij herkenning (ondanks de theoretisch hoge accuratesse die de fabrikanten opgeven) en het verspillen van resources bij het herlezen en verbeteren van de gegenereerde rapportages.

In het journal Computers and Composition uit 2008 staat een artikel van Thomas Batt en Sandip Wilson⁶¹ over het gebruik van VRS binnen het onderwijs, te weten bij het geven van feedback op door studenten ingeleverd materiaal. Er bleek dat het gebruik van VRS een handige tool kan zijn voor leraren, maar dat het door de beperkte edit- en verbeteringsvermogen niet geheel past binnen de lijnen die binnen hun onderzoek zijn uitgezet. Studenten zagen geen noemenswaardig verschil tussen commentaar op hun ingeleverde stukken gegenereerd via VRS en die via de traditionele methodes. Ook is er geen noemenswaardige tijdwinst geboekt bij het gebruik van VRS doordat de docenten hun commentaar vaker door moesten lezen om er storende spraakherkenningfouten uit te halen.

Het journal of Digital Imaging heeft in 2008 een uitgebreid artikel geweid aan het gebruik van VRS binnen de radiologie. John Pezzulo et al.⁶² hebben onderzoek gedaan naar de gebruikerstevredenheid van radiologen over rapporten die met gebruik van VRS zijn geproduceerd en tevens hebben ze gekeken naar de kwaliteit van de rapporten. Een erg opvallende conclusie was dat het 50% langer duurde om rapporten te dicteren met behulp van VRS en dat deze rapporten dan ook 24% korter waren dan rapporten gegenereerd via de traditionele transcriptie-methode. Ook hadden de VRS rapportages veel meer fouten dan de transcriptie-rapporten, zowel voor als na correctie en goedkeuring. Na het aftekenen van de rapporten door de radiologen bleek 35% van de

VRS rapporten nog steeds fouten te bevatten. Ook de kosten van transcriptie via VRS t.o.v. de traditionele methode liggen veel hoger. Uitgaande van gemiddelde salarissen voor radiologen en transcriptionisten komt de kostenverhoging neer op meer dan 6 dollar per gegenereerd rapport. De werkelijke kosten kunnen nog veel hoger liggen aangezien de radiologen ontevreden waren over de verhoogde werkdruk en de daarmee samenhangende vermoeidheid, en irritatie door fouten in de rapporten. Ondanks deze tekortkomingen spreken ze van een schatting van ongeveer 30% van de radiologieafdelingen van ziekenhuizen in de VS die van plan zijn om spraakherkenningsystemen te implementeren in de nabije toekomst.

In de technology sectie van Hospitals & Health Network magazine verschijnt een stuk geschreven door Jane Soung⁶³. Ze beschrijft dat na een succesvolle test van VRS door drie artsen van het Boston Massachusetts General Hospital en een uitgebreidere, succesvolle, pilot inmiddels 175 gebruikers op verschillende afdelingen gebruikmaken van VRS. Ondanks de kortere doorlooptijd voor de productie van rapportages kleven er zeker nog nadelen aan het gebruik van de software. Zo moeten dokters 1 oog op het beeldscherm houden tijdens het dicteren om te controleren of er niet teveel fouten gemaakt worden door de software. De last voor accurate transcriptie ligt nu op de schouder van de specialisten in plaats van de transcriptionisten en dit zorgt voor extra weerstand bij implementatie.

Karthik Natarajan et. al.⁶⁴ beschrijven op het AMIA Symposium in 2008 al dat bij een haalbaarheidsstudie voor het gebruik van VRS in een minder dan ideale omgeving voor het gebruik van geautomatiseerde transcriptieproductie. Er is gekeken naar een groep van 72 testpersonen onderverdeeld in gebruikers zonder training, minimale training en uitgebreide training en er kon geconcludeerd worden dat de met behulp van VRS gegenereerde rapportages semantisch en syntactisch onacceptabel waren.

Susan De La Paz⁶⁵ beschrijft in 2009 in het journal Learning Disability Quarterly waar gehandicapte studenten op moeten letten bij het gebruik van VRS. Een gemiddelde volwassene is in staat om 15 tot 25 woorden per minuut met een pen te schrijven, typen verhoogt de output met een factor 2, maar gemiddelde spraak gaat met een snelheid van 125 tot 160 woorden per minuut. De bestaande VRS systemen kunnen de dictatie niet met deze snelheden verwerken, maar ze zijn in staat om continue spraak 4 tot 6x sneller te transcriberen (en met relatief hoge accuratesse) dan een gemiddelde volwassene kan typen. Zij concludeert na haar onderzoek dat er meer aandacht besteed moet worden aan de VRS mogelijkheden voor gehandicapte studenten. Het is wel nog de vraag of de VRS systemen zich nog voldoende zullen ontwikkelen zodat elke student er mee kan werken zonder teveel fouten te maken. Maar de geeft aan dat de technologie zich dermate snel ontwikkelt dat er een gereede kans is dat alles studenten, dus niet alleen de gehandicapte, VRS zullen gebruiken in plaats van de traditionele methodes van dictatie zoals pen en keyboard.

Mark Hagland⁶⁶ schrijft in 2009 in het blad Healthcare Informatics het artikel "Nuance: The Spoken Word" over het gebruik van de VRS van Nuance binnen de medische wereld. Dragon NaturallySpeaking heeft volgens eigen zeggen in Noord-Amerika 100.000 gebruikers, 20 procent van de artsen in de VS maakt gebruik van hun software en de hebben 3000 ziekenhuizen als klant. Door voor verschillende afdelingen binnen het medisch specialisme verschillende pakketten aan te bieden schat Nuance dat zij 200 tot 300 miljoen dollar per jaar besparen op traditionele transcriptiekosten voor de deelnemende ziekenhuizen. In 2008 onderzocht Larry Garber, M.D., the medical director for informatics at the Fallon Clinic (Worcester, Massachusetts) of deze claims wel terecht waren. Ze hadden namelijk aan aantal jaar daarvoor getest met een eerdere versie van de software van Nuance en waren ontevreden over het behaalde resultaat. Een week lang hebben 10 artsen gevolgd tijdens hun normale werkwijze. Vervolgens zijn dezelfde artsen met Dragon NaturallySpeaking gaan werken. Het resultaat was een afname van 95% in de hoeveelheid aangeboden dictaties en transcripties. Tevens nam de

verkrijgbaarheid van de aantekeningen van de artsen dramatisch toe. In het verleden werden de dictaties in het buitenland getranscribeerd. Gemiddeld duurde het twee dagen voordat de transcripties terugwaren in het ziekenhuis en nog twee dagen voordat het rapport goedgekeurd werd door de specialist. Deze tijd werd door het gebruik van VRS gereduceerd tot 46 minuten. De Fallon Clinic is daarna snel overgegaan op het bedrijfswijde gebruik van de software.

Ook in 2009 verscheen op de website geek.com, een website voor mensen die gebruik maken van nieuwe technologieën, een gebruikersreview van de VRS van Nuance door Jim Hatley⁶⁷. Het viel hem op dat hij zijn spraakmethode enigszins aanpast ten opzichte van een gesprek dat hij zou voeren met een echt persoon. Hij legde meer nadruk op de uitspraak en de articulatie. Na een korte inwerktijd werkte de software al naar behoren. Hij moest nog steeds de gegenereerde tekst nalezen en fouten corrigeren, maar het dicteren voelde een stuk efficiënter aan dan het zelf typen van tekst. Zijn conclusie was dat dit product niet geschikt zou zijn voor elke gebruiker, maar dat het wel een glimp laat zien van de toekomst. Hij verwacht dat door de snelheid van de ontwikkeling van de technologie die nu alleen in Hollywoodfilms te zien is, binnen een aantal jaar VRS beschikbaar zal zijn voor het gewone publiek.

In 2009 verschijnt in het *Journal Clinical Radiology* een artikel van N.Goyal et al.⁶⁸ over bijkomende (ergonomische) voordelen van het gebruik van VRS binnen de radiologie. Hij signaleert een afname van fysieke klachten (vooral in de pols, onder- en bovenarm, schouder en nek) bij radiologen die gebruik maken van VRS. Omdat een radioloog veel tijd kwijt is met de rapportages kan VRS een positief effect hebben op de gezondheid. Mogelijke gevallen van RSI kunnen door het gebruik van VRS voorkomen worden. Hij concludeert dat het volle potentieel van nieuwe technologieën, met toenames in efficiëntie, effectiviteit en algemene werkzaamheden bij het gebruik van VRS nog niet volledig benut wordt.

5 Antwoord op de onderzoeksvragen

In dit hoofdstuk zal een antwoord gegeven worden op de onderzoeksvragen die in hoofdstuk 2 zijn beschreven. Ook zal er, waar mogelijk, ingegaan worden op de bevindingen uit de literatuur daar waar relevante artikelen gevonden zijn.

5.1 Wetenschappelijke visie

Vanuit de literatuur is duidelijk dat wetenschappers uit verschillende disciplines vinden dat er een gouden toekomst voor Voice Recognition Software (VRS) is weggelegd. Doordat mensen sneller kunnen spreken dan typen en de techniek achter VRS en die van de IT in het algemeen inmiddels dermate ver gevorderd is dat deze ook geen belemmering meer vormt, zou men verwachten dat VRS al uitgebreid gebruikt zou worden in omgevingen waar veel gebruik gemaakt wordt van dictatie, het genereren van rapportages en de verwerking van grote hoeveelheden tekst. Daar het gebruik van e-mail binnen bedrijven gemeengoed is geworden en dit een grote aanslag pleegt op de tijd van werknemers zou men verwachten dat ook hier gebruikt gemaakt zal worden van VRS. Dit blijkt echter niet zo te zijn. Er blijkt dat alleen in het veld van de radiologie veel gebruik gemaakt wordt van deze technologie. Hoe komt dit?

Het is duidelijk dat de fabrikanten van VRS claimen software te ontwikkelen die bijna foutloos werkt en erg gebruiksvriendelijk is. Uit de bestudeerde literatuur blijkt dit in de praktijk echter tegen te vallen. Gebruikers moeten langdurig getraind worden voordat ze de software goed kunnen gebruiken en zelfs dan maakt de software nog veel fouten. Deze fouten moeten vervolgens, hetzij door de gebruiker zelf, hetzij door een corrector, verbeterd worden en dit zorgt ervoor dat een groot gedeelte van de tijd die gewonnen is door gebruik te maken van spraakherkenning weer verloren gaat. Ook verschuift de druk van de (foutloze) rapportering van de relatief goedkope administratieve krachten naar die van de verantwoordelijken, meestal managers en artsen. De productiviteit van deze dure werknemers neemt af waardoor de winst die in het administratieve proces wordt behaald weer tenietgedaan wordt.

Tevens valt op dat er in de literatuur nergens een praktijkvoorbeeld beschreven wordt van het gebruik van VRS bij de afhandeling van e-mail. Dat het mailen tegenwoordig een belangrijk onderdeel van de werkzaamheden van werknemers uitmaakt is evident, maar hulpmiddelen om het omgaan met mail te vergemakkelijken worden niet beschreven. Vooral de procedures die gebruikt moeten worden om de afhandeling van e-mail te standaardiseren worden beschreven, maar tools om de afhandeling te versnellen zijn er niet, of wellicht nog in ontwikkeling. Binnen de wetenschap leeft het idee dat men de werknemers eerder dient op te voeden over het gebruik van e-mail, de zogenaamde netiquette, om de druk te verminderen dan dat andere technologieën de werknemers bij het verlichten van de (e-mail) werkdruk kunnen helpen.

Dat de werkdruk door de enorme toename van het elektronische dataverkeer extreem toeneemt is in de literatuur vaak beschreven. Dat overloadgevoelens leven onder een gedeelte van de werknemers wordt door een flink aantal publicaties bevestigd. Maar naar de opmerking van Ron Weber², dat spraakherkenning wellicht een oplossing kan zijn om overloadgevoelens weg te nemen, is nooit serieus onderzoek gedaan. Onderzoeken die zijn uitgevoerd over het functioneren van VRS komen over het algemeen uit die disciplines die omgaan met gehandicapten en computergebruik en uit de medische wereld. De informatietechnologie heeft het potentieel van deze technologie wel onderkend, maar doet verder weinig onderzoek naar het gebruik ervan. Dit zou een gemiste kans kunnen zijn.

5.2 VRS en e-mail

Zoals hiervoor beschreven zijn er geen duidelijke publicaties te vinden over het gebruik van VRS bij de afhandeling van e-mail. Het feit dat door een aantal wetenschappers

wordt beweerd dat voor bijvoorbeeld privé-communicatie en het maken van privé-aantekeningen de technologie, ondanks de fouten, erg nuttig kan zijn is blijkbaar nog niet in de wereld van de IT doorgedrongen. Men focust teveel op het kostenaspect en de fouten die gemaakt worden door de software. Dit hoeft bij dictaties voor eigen gebruik of simpele afhandelingen via e-mail niet nadelig te zijn. Men ziet de technologie momenteel meer als leuk speeltje dan dat er serieus gebruik van gemaakt wordt. De radiologie is gelukkig een positieve uitzondering op deze conclusie.

Het verdient dus de aanbeveling om in verder onderzoek te testen of de technologie, die zich sneller en sneller door ontwikkelt, in de toekomst gebruikt kan worden voor meer. Voor werkzaamheden waar een grote nauwkeurigheid niet vereist is, maar wel veel gebruik gemaakt wordt van transcripties en tekstverwerken kan VRS een grote kostenbesparing opleveren. Dit zou zelfs nu al plaats kunnen vinden, gezien de huidige staat van de technologie. Te denken valt aan het schriftelijk behandelen van uitgebreide vragen aan fabrikanten of dienstverleners, waar geen gebruik gemaakt kan worden van gestandaardiseerde antwoordformulieren vanwege de variatie in de vragen. Daar waar een persoon nu veel tijd kwijt is aan het verwerken van tekst kan de software deze taak sneller doen en dus vaak ook goedkoper.

Voor het gebruik van VRS in situaties waar de nauwkeurigheid van de gegenereerde tekst erg belangrijk is zal de toekomstige ontwikkeling van de software afgewacht moeten worden. Daar waar een fout in gegenereerde tekst binnen e-mails grote consequenties kan hebben (bijvoorbeeld binnen de gezondheidszorg, juridische correspondentie etcetera) is het nu nog niet raadzaam om volledig te vertrouwen op deze technologie. In de toekomst zal dit wellicht nog veranderen.

5.3 Toekomstperspectief VRS

De laatste vraag die beantwoord moet worden is of het gebruik van VRS een toekomst heeft binnen een bedrijfsomgeving. Het feit dat binnen de radiologie de rapportagefunctie meer en meer overgenomen wordt door VRS lijkt deze vraag positief te beantwoorden. Wel dient hierbij opgemerkt te worden dat een controle binnen deze discipline altijd ingebouwd moet zijn. Als er een fout gemaakt wordt door de software en dit kan invloed hebben op de gezondheid van patiënten dan moet men er zeker van zijn dat dit niet plaats kan vinden. Het is niet voor niets dat het leger op laag niveau gebruik wil maken van deze technologie, maar voor onderdelen waar een kleine vergissing grote consequenties kan hebben lijkt deze technologie nog niet goed genoeg.

Gebruikers moeten goed getraind worden voor het gebruik van de software en men zou de VRS alleen moeten gebruiken als deze beschikbaar is in zijn/haar moedertaal. Ook moet er een goede omgeving beschikbaar zijn om gebruik te maken van deze software. Gezien de grote invloed die ruis en omgevingsgeluid kunnen hebben of het functioneren is een afgesloten ruimte en het gebruik van goede hardware (zowel computers met voldoende rekencapaciteit als kwalitatief hoogwaardige microfoons) van belang.

Gebruikers moeten achter het gebruik van VRS staan en niet het gevoel hebben gedwongen te worden tot het gebruik van deze technologie. Als men om kan gaan met VR-systemen kan dit helpen om een hoop (administratieve) taken sneller uit te voeren. Overloadgevoelens die hier uit voortkomen kunnen dan worden weggenomen. Ook kan een verbetering van de productiviteit optreden doordat men sneller tekst kan produceren met behulp van VRS dan met de traditionele methodes als typen. Ook kan er bespaard worden op de transcriptiekosten, doordat gebruikers hun rapporten zelf kunnen produceren en ook zal de beschikbaarheid van de rapporten versneld worden, doordat er niet een extra tussenliggende transcriptiestap nodig is.

Of in de toekomst meer disciplines gebruik gaan maken van VRS moet nog blijken. Het potentieel is onderkend, maar er moet nog uitgebreid onderzoek gedaan worden naar de kwaliteit van de verschillende VRS-systemen en het gebruikersgemak. Ook is er nog af te wachten of VRS wellicht deel gaat uitmaken van de hardware van de computers van

de komende jaren. Hoe meer VR geïntegreerd gaat worden in de computer van de toekomst, hoe groter de kans dat deze technologie gemeengoed wordt. Als al deze puzzelstukjes in de toekomst op zijn plaats vallen kan VRS een belangrijk onderdeel van onze interactie met computers uit gaan maken. De trainingsbehoefte van personeel, die nu nog vereist is bij het gebruik van VRS, is in de toekomst zeer waarschijnlijk een stuk minder. Men groeit op met het gebruik van VRS en leert van jongs af aan dus met deze technologie om te gaan.

Het feit dat binnen de radiologie verwacht wordt dat in de komende 5 tot 10 jaar het overgrote deel van de rapporten via VRS gegenereerd zal worden^{33,35} geeft aan dat het bedrijfsleven langzaam beseft welk potentieel het gebruik van VRS kan hebben in de toekomst. Het is derhalve een kwestie van tijd voordat andere disciplines binnen het bedrijfsleven dit potentieel ook zullen onderkennen en zullen onderzoeken of het gebruik van VRS binnen hun bedrijfstak ook van nut kan zijn.

In onderstaande tabel staan de voor- en nadelen van het toekomstige gebruik van VRS bij verschillende toepassingen nog kort uitgelegd.

Argument	Omschrijving	Voordelen gebruik VRS	Nadelen gebruik VRS
Transcriptie via VRS	Traditioneel gebruik van dictafoon. In plaats van het uitwerken van de tekst door een transcriptionist gebruik maken van VRS gekoppeld aan de dictafoon.	1) Kostenbesparing op transcriptiekosten 2) Minder kans op RSI gebruiker	Door de fouten die door de VRS gemaakt worden zal de tekst niet 100% overeenkomen met de dictatie. Bij vereiste grote nauwkeurigheid geen optie. Anders afhankelijk van tolerantie foutenmarge.
Realtime gebruik VRS	Gesproken woorden tegen de computer worden realtime omgezet naar tekst.	Gebruiker spreekt sneller dan dat hij kan typen en dus bespaart de gebruiker tijd.	1) Gebruiker dient in een stille omgeving te dicteren om de fouten door omgevingsgeluid te minimaliseren. 2) Grotere netwerkbelasting en toename hardware-eisen bedrijven door VRS
VRS en e-mail	Gebruik VRS bij afhandeling e-mail.	1) Afname overloadgevoelens bij overmatig e-mail gebruik. 2) Gebruiker heeft handen vrij voor overige taken 3) Tekstgeneratie met behulp van VRS gaat sneller dan typen	1) Omgevingsgeluid kan invloed hebben op de foutenmarge. 2) E-mail afhandeling betreft vaak korte berichten waardoor tijdwinst minder invloed heeft. 3) Gebruikers hebben additionele training nodig bij gebruik e-mail en VRS
VRS in het algemeen in de toekomst	Technologie ontwikkelt zich in de toekomst verder en is meer gemeengoed bij computergebruikers.	1) Afnemende foutenmarge 2) Afnemende kosten aanschaf en onderhoud VRS pakketten 3) Afname kans op RSI bij verwerking grote hoeveelheden tekst 4) Benutting potentieel arbeidsmarkt van manueel gehandicapte medewerkers	1) Gebruiker voelt zich meer machine dan mens 2) Kans op stemproblemen door continu gebruik van spraak 3) Duurbetaalde gebruikers zijn meer tijd kwijt dan bij delegeren tekstverwerking en dus nemen de overall kosten toe 4) Grotere trainingsbehoefte werknemers omgang met VRS

6 Conclusie

In de serie Star Trek, een Sci-Fi serie uit de jaren zestig, wordt al gebruik gemaakt van stemmen om te communiceren met computers. De technologie is inmiddels dermate ver ontwikkeld, dat spraakherkenning geen science fiction meer is. Waarom wordt er dan toch zo weinig gebruik van gemaakt in het bedrijfsleven en is er zo weinig informatie over deze technologie te vinden?

Als sinds 1993 is bij verschillende wetenschappers uit de disciplines van de communicatie, informatietechnologie en in een later stadium ook vanuit de medicijnen duidelijk dat spraakherkenning in de toekomst een belangrijke functie kan vervullen. Er wordt sinds de jaren 90 een gouden toekomst voor VRS voorspeld, maar helaas is deze toekomst tot op heden nog niet waargemaakt. Wel zijn er verschillende onderzoeken uitgevoerd (Goette²⁷, Mangold²⁹, Wald en Bain⁵²), om te bepalen aan welke voorwaarden VRS moet voldoen om uit te groeien tot een volwaardig onderdeel van het digitale desktop-pakket.

Dat training een belangrijk onderdeel van de toekomst van deze technologie gaat bepalen is uit de literatuur goed af te leiden. Hoe beter men om weet te gaan met de technologie, hoe makkelijker in de toekomst onze interacties met computers zal verlopen. Voordeel van veel in aanraking komen met moderne technologie is dat men makkelijk went aan nieuwe ontwikkelingen op dat gebied. Hoe meer mensen werken met nieuwe technologie, hoe makkelijker het is voor nieuwe gebruikers om te leren van anderen en zo ook snel om te gaan met de technologie. Dit is vergelijkbaar met het feit dat één ervaren radioloog in Finland het voor elkaar gekregen zijn om de hele afdeling te onderwijzen in het systeem van VRS⁵⁹. De na-oorlogse generatie heeft inmiddels geleerd om te werken met computers via traditionele interactie-methodes als keyboard en moderne methodes als touch-screens. Waar enkele tientallen jaren geleden weerstand bestond tegen deze digitale revolutie is deze nu algemeen geaccepteerd en zo zal het wellicht ook verlopen met het gebruik van VRS.

Het gevoel van overload dat bij managers momenteel leeft door de grote hoeveelheden binnenkomende e-mail zal wellicht afnemen als men meer gestandaardiseerd en gestructureerd gebruik maakt van dit communicatiemiddel. Dit is al gesignaleerd door Jeanne Allart¹², Jenni Ingham¹⁵, David Dawley¹⁶, McKeen et al.¹⁷, Gil Thomas et al.⁵ en Ashish Gupta et al.²⁰. Gezien deze grote hoeveelheid wetenschappelijk bewijs dat verkeerd gebruik van e-mail overloadgevoelens veroorzaakt is het aan te raden dat bedrijven waar dit speelt eens goed kijken welke e-mailprotocollen er in gebruik zijn. Mochten deze binnen hun bedrijf niet bestaan dan zou het opstellen van een interne 'netiquette' wellicht bij kunnen dragen aan de afname van de overloadgevoelens.

Tegenwoordig schrikt spraakherkenning de meeste gebruikers niet meer af en komt men al op jonge leeftijd in aanraking met deze moderne technologie. Via mobiele telefoons en smartphones, die tegenwoordig bestempeld kunnen worden als mini-computers, worden de barrières tussen mensen en machines steeds verder afgebroken. VRS zit al standaard in de nieuwste modellen telefoons en tablets en in de nabije toekomst zal dit wellicht ook doordringen in het algemene computergebruik. Mocht het deel gaan uitmaken van een standaard computer, zoals Linda Paulson beschrijft⁵¹, dan zal de acceptatie een kwestie van tijd zijn. Of onze kinderen in de toekomst zullen praten tegen hun computers en geen gebruik meer hoeven te maken van keyboard en muis zal echter nog moeten blijken. Bijkomend voordeel van deze verandering is dat zij minder last zullen hebben van moderne computerziektes als RSI (N. Goyal⁶⁸ maakt al melding van dit voordeel in 2009). Dat zij makkelijker zullen omgaan met deze technologie en het minder als toekomstmuziek zullen bestempelen is echter een zekerheid.

Referenties

- a) http://e-mail.about.com/od/e-mailtrivia/f/how_many_e-mail.htm
1. R.B. Schwarz, M.C. Russo, How to quickly find articles in the top IS Journals, 2004
 2. Ron Weber, The Grim Reaper: The Curse of E-Mail, 2004
 3. Steve Whittaker, E-mail overload: exploring personal information management of e-mail, 1996
 4. McShane & Von Glinow, Organizational Behaviour, 2005
 5. Gail Fann Thomas et al., Reconceptualizing e-mail overload, 2006
 6. Peter J. Denning- Electronic Junk, 1982
 7. S. Hiltz, M. Turoff – Structuring Computer-Mediated Communication Systems to avoid Information Overload, 1985
 8. Janet Fulk - Electronic Communication and Changing Organizational Forms, 1995
 9. Charles Arthur - Information overload makes managers ill, 1996
 10. V. Houlder - Failing to get the message, 1997
 11. G. Ruggeri Stevens - A qualitative study and model of the use of e-mail in organisations, 2000
 12. Jeanne L. Allert, A 12-Step (or so) Program for Information Junkies, 2001
 13. Ralph Adam - Is e-mail addictive, 2002
 14. Jenni Ingham - E-mail overload in the UK workplace, 2003
 15. Mallon and Oppenheim, Style used in Electronic Mail, 2002, p. 15)
 16. David D. Dawley, User perceptions of e-mail at work, 2003
 17. James D. McKeen et al. - Developments in Practice XIII Electronic Communications - Strategies for Coping with the Deluge, 2004
 18. Schuff et al., A multi-attribute, multi-weight clustering approach to managing e-mail overload, 2005
 19. Belotti et al., Quality versus quantity E-mail-centric task management and its relation with overload, 2005
 20. Gupta et al. - E-Mail Management - A Techno-Managerial Research Perspective, 2006
 21. Spoelstra - The Problem of E-mail Overload: A Public Sector Case Study, 2007
 22. Span - The Problem of E-mail Overload. A Private Sector Case Study, 2007
 23. J. Markowitz, Using Speech Recognition, 1996
 24. Farzad Ehsani, Eva Knodt, Speech Technology in Computer-aided Language Learning: strengths and limitations of a new call paradigm, 1998
 25. D. Coniam, Voice recognition software accuracy with second language speakers of English, 1999
 26. G. Alwang, Speech Recognition, 1999
 27. Tanya Goette, Keys to the adoption and use of voice recognition technology in organizations, 2000
 28. M.J. Haxer et al., Use of Speech Recognition Software: A Vocal Endurance Test for the New Millennium, 2001
 29. Helmut Mangold, Speech Technology in Reality – Applications, Their Challenges and Solutions, 2001
 30. Charles Lowe, Speech Recognition: Sci-Fi or Composition, 2001
 31. Sarah Harris, Alan Biermann, Mouse Selection Versus Voice Selection of Menu Items, 2002
 32. Deng et al., Speech and Language Processing for Multimodal Human-Computer Interaction, 2002
 33. Amit Mehta, MD, and Theresa C. McCloud MD, Voice Recognition, 2003
 34. K. Larson en D. Mowatt, Speech Error Correction: The Story of the Alternates List, 2003
 35. M.Grasso, The Long-Term Adoption of Speech Recognition in Medical Applications, 2003
 36. James Hartley, Speaking versus typing: a case-study of the effects of using voice-recognition software on academic correspondence, 2003
 37. JM Rieger, The effect of automatic speech recognition systems on speaking workload and task efficiency, 2003
 38. Robert Issenman, Use of Voice Recognition Software in an Outpatient Pediatric Specialty Practice, 2004
 39. Bryan Bergeron, Technology in Your Practice - Voice Recognition and Medical Transcription, 2004
 40. Atif Zafar et al., A simple error classification system for understanding sources of error in automatic speech recognition and human transcription, 2004
 41. Nick van Terheyden M.D., Is Speech Recognition the Holy Grail, 2005
 42. Feng et al., How productivity improves in hands-free continuous dictation tasks: lessons learned from a longitudinal study, 2005

43. DeRosier et al., Speech recognition software as an assistive device: A pilot study of user satisfaction and psychosocial impact, 2005
44. Keith S. White, Speech recognition implementation in radiology, 2005
45. Stephen Manes , Voice Recognition Ready for Prime Time, 2005
46. Richard L. Morin en Steve G. Langer, Speech Recognition System Evaluation, 2005
47. Sixto Ortiz Jr., The state of voice recognition: end users offer their perspectives on its use in the business market, 2005
48. Julie Park and A. Echo Zeanah, An evaluation of voice recognition software for use in interview-based research: a research note, 2005
49. Fang Chen, Designing human interface in Speech Technology, 2006
50. Giovanni Cannata - Phymail Box, an information appliance that checks and prints only important e-mails, 2006
51. Paulson, Speech Recognition Moves from Software to Hardware, 2006
52. Mike Wald, Keith Bain, Universal access to communication and learning: the role of automatic speech recognition, 2007
53. John Binder, Speak, don't type, 2007
54. V. Keegan, Has voice recognition finally come of age?, 2007
55. Laura Cantrall Cordio, Speak and Ye Shall Be Heard, 2007
56. D.J. Quint, Voice Recognition: Ready for Prime Time?, 2007
57. Wald et al., Correcting automatic speech recognition captioning errors in realtime, 2008
58. András Kocsor ·Gábor Gosztolya , The use of speed-up techniques for a speech recognizer system, 2008
59. Mika P. Koivikko et al., Improvement of Report Workflow and Productivity Using Speech Recognition—A Follow-up Study, 2008
60. Kimberly Voll , Improving the utility of Speech Recognition through Error detection, 2008
61. Batt, Sandip, A Study of Voice-Recognition Software as a Tool for Teacher Response, 2008
62. John Pezzulo et al., Voice Recognition Dictation: Radiologist as Transcriptionist, 2008
63. J. Soung, Speech Recognition Software Wins Converts, 2008
64. Karthik Natarajan et al., Feasibility Study of Speech Recognition for Gathering Information Needs, 2008
65. Susan De La Paz, Composing via Dictation and Speech Recognition Systems: Compensatory Technology for Students with Learning Disabilities, 2009
66. Mark Hagland, Nuance: The Spoken Word, 2009
67. Jim Hatley, Dragon NaturallySpeaking Voice Recognition Software review online, 2009
68. N.Goyal, Ergonomics in radiology, 2009