

Spraakgestuurde applicaties staan toe dat via een mobiele of vaste telefoon altijd en overal informatie-uitwisseling kan plaatsvinden met databanken, applicaties en bedrijfsinformatiesystemen. En dat computers met spraakherkenning en -synthese menselijke activiteiten kunnen vervangen. Tot voor kort ging dit gepaard met hoge investeringen en ingewikkelde koppelingen aan andere applicaties/systemen. De huidige technologische ontwikkelingen brengen spraakgestuurde dienstverlening steeds dichterbij, wellicht ook voor uw organisatie...

# Look who's talking...

Dirkjan van Groeningen

De ontdekking dat spraak kan worden omgezet in elektromagnetische deeltjes heeft de ontwikkeling van spraaktechnologie een enorme impuls gegeven. In feite is spraaktechnologie onder te verdelen in drie disciplines: spraakherkenning, spraak-synthese en sprekerherkenning. Door spraakherkenning kan apparatuur bediend worden via gesproken commando's en ontstaat een alternatief voor het handmatig invoeren van gegevens. Door spraak-synthese kunnen computers spraak reproduceren, dialogen voeren en uiteenlopende vragen beantwoorden over bijvoorbeeld treinregelingen, toeristische attracties en files. Een computer herkent en analyseert de vraag, vervolgens geeft een gesynthetiseerde stem antwoord. Sprekerherkenning is interessant voor identificatie en verificatie van personen. Naast informatiediensten kunnen spraakgestuurde computersystemen bijvoorbeeld worden ingezet voor interactieve leermethodes, dicteesystemen, en tekst-naar-spraakomzetsystemen voor visueel gehandicapten.

Bekend is vooral het gebruik van spraaktechnologie voor afhandeling van delen van het telefoonverkeer. Dit is veelal interactive voice response (IVR) in combinatie met sturing door telefoontoetsen. Bij voice mail bijvoorbeeld wordt vooral gebruik gemaakt van de druktoetsen op de telefoon: "kies 1 om uw voicemail af te luisteren". Door combinatie van IVR met spraakherkenning wordt een automatische dialoog niet alleen efficiënter, maar ook flexibeler en klantvriendelijker. Standaardvragen die geen persoonlijke interactie vereisen kunnen direct afgehandeld worden. Daardoor krijgen medewerkers de ruimte om extra aandacht te besteden aan complexe of nieuwe vragen en problemen.

## *Anytime, anyplace, anywhere*

Spraaktechnologie kan overal worden ingezet waar gebruikers en systemen boodschappen uitwisselen. De combinatie van webtechnologieën, spraakgestuurde diensten en telefonie-mogelijkheden inspireert tot interessante zakelijke scenario's. Denk aan een accountmanager die onderweg naar de klant via een beveiligd, spraakgestuurd systeem de laatste orderstatus van deze klant opvraagt. Een live 'track & trace' applicatie van bestellingen, of telefonische toegang tot e-mail als er geen internetverbinding is. Ook is spraakherkenning inzetbaar in magazijnen, bij in- en uitslaan van goederen, verplaatsen van pallets en sorteren van goederen en ontvangstcontrole. Voor het uitwisselen van opdrachten en bevestigingen hoeft dan geen scanner meer ter hand genomen te worden. Uitwisseling van boodschappen gebeurt grotendeels in combinatie met andere acties, wat tijd bespaart.

## *Nieuwe ontwikkelingen*

Defacto standaard voor het koppelen van data en telefonie met spraak als gebruikersinterface is de 'Voice Extensible Markup Language' ofwel Voice XML. XML is een taal waarmee op gestructureerde wijze informatie vastgelegd kan worden met behulp van 'tags'. De tags geven de semantiek (betekenis) weer van de woorden die tussen de tags staan. Voice XML is een schema waarin de opmaak van een dialoog tussen content, meestal in de vorm van een webpagina, en beller kan worden vormgegeven met mogelijkheden voor gebruik van gesproken commando's, vooropgenomen spraak, muziek en spraak-synthese. Door dergelijke nieuwe ontwikkelingen ontstaat een onafhankelijke integratie met de business logica.

## Pop-Eye en Olive

Toepassingen die gezien de toegenomen hoeveelheid videobestanden interessant zijn, zijn Pop-Eye en Olive, mede-ontwikkeld door specialisten van de Universiteit Twente. Pop-Eye is een systeem dat de ondertitels van audiovisueel materiaal indexeert en gedeeltelijk vertaalt. Olive indexeert aan de hand van spraaktechnologie geluidsbanden. Met behulp van deze systemen kunnen beeld- en geluidsarchieven worden geraadpleegd aan de hand van zowel geschreven als gesproken sleutelwoorden. Hierdoor is het nu bijvoorbeeld mogelijk om naar specifieke fragmenten te zoeken van geluids- en beeldbestanden van de bekende Nederlandse schrijver W.F. Hermans. Volgende stap is nu de beschikbaarheid van dit audiovisuele archief via [www.willemfrederikhermans.nl](http://www.willemfrederikhermans.nl).

### Generieke toepassingen

- Automatisch aanvragen van brochures en overige documenten
- Systemen ter beantwoording van FAQ's, veelgestelde vragen
- Het routeren en classificeren van gesprekken met behulp van spraakherkenning, intelligente wachtrijen, opvang piekverkeer
- Voicemail/call recording
- Identificatie van bellers op basis van klantnummer of postcode en huisnummer, toegang tot gepersonaliseerde menus, opslaan van naw-gegevens

### Media & professional services

- Het aanvragen van (proef)abbonementen
- Opvang bezorgklachten
- Doorgeven vakantieopgave
- Doorgeven adreswijziging
- Automatische enquêtes, marktonderzoeken, verkiezingen en pellingen
- Verwerking aanvragen n.a.v. advertenties
- E-mail reading, het voorlezen van actueel nieuws

### Branche-specifieke toepassingen

#### Chemie & pharma

- Informatie over afdelingen/divisies
- Klachtenopvang
- Informatie over spaar-/loyalty-acties
- Persoonsgebonden informatie (spaarsaldo, benzineverbruik etc.)
- Ordermogelijkheid voor klant
- Productondersteuning

#### Manufacturing

- Eerstelijns opvang voor klanten buiten kantooruren
- Productinformatie
- Informatie over productstatus van een klant
- Ordermogelijkheid voor klant
- Klanteninformatie voor accountmanagers onderweg
- Helpdeskondersteuning
- Beantwoording van veelgestelde vragen
- Automatisch aanvragen van offertes/brochures

#### Consumer products & retail

- Eerstelijns opvang voor klanten buiten kantooruren
- Productinformatie
- Informatie over productstatus van een klant (tracking & tracing)
- Enquêtes onder klanten
- Klanteninformatie voor accountmanagers onderweg
- Helpdesk-ondersteuning
- Beantwoording van veelgestelde vragen
- Aanvragen van offertes/brochures

De spraakinterface functioneert in feite als een onafhankelijke laag die verschillende programmeertalen kan ondersteunen. Hierdoor worden spraakherkenningapplicaties steeds beter beheersbaar op het gebied van schaalbaarheid, compatibiliteit, hergebruik, veiligheid, onderhoud en installatie.

In juni 2004 zette Microsoft een belangrijke stap op deze markt met de introductie van Speech Server: een webgebaseerd, geïntegreerd platform voor spraakgestuurde IVR- en internet-applicaties. Microsoft Speech Server stelt ontwikkelaars in staat om hun applicaties uit te breiden met ondersteuning voor spraak. Zo zijn snel applicaties in te voeren voor gebruik vanaf de (mobiele) telefoon, pocket pc, tablet pc en andere apparaten. De verkiezing van dit product tot 'bedrijfs-

infrastructuurproduct van het jaar 2005' door onderzoeksbureau Frost & Sullivan is een belangrijke indicatie dat spraaktechnologie de kinderschoenen ontgroeit.

### Acceptatie

Net als bij andere nieuwe technieken in het verleden roept spraakherkenning bij velen wantrouwen op: computers die luisteren, praten en handelen. Het tij begint echter langzaam te keren, wellicht door de succesvolle business cases die er al zijn bij financiële dienstverleners, ziekenhuizen, onderzoeksbureaus en productiebedrijven. Het kwaliteitsniveau van spraakherkenningssoftware is nu zodanig, dat de toepassing ervan gebruikers positief zal verrassen, mits de techniek zorgvuldig en weloverwogen wordt ingezet. ■