

Uitgeschreven tekst bij hoofdstuk 1 Opdracht 5

Je zegt minder dan je denkt

<https://vimeo.com/34781438>

Let op! Om het fragment te bekijken moet je een (gratis) account bij Vimeo aanmaken.

Duur fragment: 5:12 minuten

Datum uitzending: 2011

Inleiding:

‘Tuuk’, ‘daak’ en ‘fosmij’, woorden die je misschien niet bekend voorkomen, maar we gebruiken ze iedere dag. Zo spreken we ‘natuurlijk’, ‘dadelijk’ en ‘volgens mij’ uit als we er niet bij nadenken. Er is een groot verschil tussen hoe we taal netjes uitspreken en hoe woorden klinken in alledaagse gesprekken. En toch begrijpen we elkaar prima. Hoe doen we dat? Taalkundige Mirjam Ernestus bouwt wiskundige modellen om te begrijpen hoe onze hersenen taal verwerken.

Mirjam Ernestus:

Voor mijn promotieonderzoek wilde ik weten hoe kleine woordjes zoals ‘t’, ‘m’ en ‘er’ uitgesproken worden. Maar toen kwam ik erachter dat ik dat eigenlijk helemaal zelf niet wist en daarom besloot ik heel veel spontane conversaties op te nemen – en dat is een corpus – om uit te zoeken hoe die woorden nou precies uitgesproken worden. Maar toen ik eenmaal dat corpus had en in dat corpus ging kijken, toen bleek dat ik van heel veel woorden helemaal niet wist hoe ze werkelijk uitgesproken werden, want heel veel woorden komen dus veel voor ... korter voor dan je zou verwachten.

Binnen de taalkunde, de psycholinguïstiek, en ook alle andere vakgebieden die zich bezighouden met gesproken taal is het heel erg de gewoonte om te kijken naar hoe woorden in formele situaties uitgesproken worden. Dus hoe ze heel erg netjes uitgesproken worden. Maar ik doe juist onderzoek naar hoe woorden in spontane conversaties uitgesproken worden, waar ze dus heel vaak gereduceerd worden. En dat gebeurt tot nu toe nog steeds heel erg weinig.

We willen graag weten hoe mensen die gereduceerde uitspraakvarianten nu precies produceren en hoe luisteraars die vormen verstaan. Dus wat we willen weten is wat er precies in de hersenen gebeurt als we die vormen uitspreken of begrijpen.

We gebruiken drie verschillende methodes om dit onderzoek te doen. Ten eerste kijken we naar corpora. Dat zijn dus die verzamelingen opgenomen conversaties. En daarin kijken we welke uitspraakvarianten er nou precies voorkomen. Want ‘tuuk’ is bijvoorbeeld maar één uitspraakvariant van ‘natuurlijk’, maar in feite alles wat mogelijk is tussen ‘natuurlijk’ en ‘tuuk’ vind je in spontane spraak. Dus we kijken welke uitspraakvarianten er voorkomen en wanneer ze voorkomen, dus hè, in welke condities precies.

Daarnaast gebruiken we experimenten, en in die experimenten proberen we ook te kijken wanneer bepaalde uitspraakvarianten voorkomen. Dus we proberen ervoor te zorgen dat proefpersonen onder bepaalde condities die varianten gaan uitspreken. Maar we gebruiken experimenten vooral om te onderzoeken hoe luisteraars die uitspraakvarianten begrijpen. Dus bijvoorbeeld

hoelang het duurt voordat ze een variant begrepen hebben en of ze die variant ook nog goed kunnen begrijpen als ze geen context krijgen, en alleen maar het woord in isolatie.

En ten derde gebruiken we wiskundige modellen. Het idee is dat we theorieën bouwen op basis van de data die we krijgen uit de corpora en de experimenten. En die theorieën die stoppen we dan in de computer, daar maken we een computermodel van. En daarmee kunnen we goed testen of die theorie ook werkelijk de goede voorspellingen doet. En daarnaast kunnen we met zo'n computermodel ook nieuwe voorspellingen doen en die kunnen we dan weer testen op basis van de corpora en experimenten.

We kijken niet alleen naar de varianten in één bepaalde taal, en ook niet alleen naar hoe luisteraars van één bepaalde taal of ... of sprekers van één bepaalde taal die varianten verwerken. En de reden daarvoor is dat we geïnteresseerd zijn in hoe mensen in het algemeen met dat soort uitspraakvarianten omgaan en daarom moeten we naar verschillende talen kijken. En wij kijken naar een aantal Europese talen: het Nederlands, het Duits, het Engels, het Frans en het Spaans.

We willen uiteindelijk begrijpen hoe sprekers en luisteraars met spontane conversaties omgaan. En een van de belangrijke vragen daarbij is wat er nu precies in het mentale lexicon is opgeslagen. Vroeger dacht men dat alleen maar de basiswoorden in dat mentale lexicon stonden, zoals bijvoorbeeld 'hond' en 'kat' en dat was het dan, en niet 'honden' en 'katten'. Maar tegenwoordig denkt men daar heel anders over. Dat kun je je ook wel voorstellen, want vroeger hadden we natuurlijk computers met heel weinig geheugenruimte, dus het idee was dan ook dat ook in onze geheugenruimte er ook alleen maar weinig plaats was voor woorden. Maar tegenwoordig zijn computers heel erg groot, kun je heel erg veel opslaan, dus nu hebben wij ook steeds meer het idee dat er veel meer in dat mentale lexicon opgeslagen zou kunnen zitten. En een van de vragen waar wij ons bezig mee houden, is: welke uitspraakvarianten zijn er nu allemaal opgeslagen in het mentale lexicon en hoe.

Ons onderzoek richt zich vooral op de gereduceerde uitspraakvarianten die we in spontane conversaties tegenkomen, maar spontane conversaties verschillen in nog veel meer opzichten van formele, nette spraak. Bijvoorbeeld, in spontane conversaties maken mensen spreekfouten. En ze spreken dan weer eens heel snel, dan weer heel langzaam, ze aarzelen, ze herhalen woorden, ze beginnen met een zin, stoppen op de helft, beginnen dan met de volgende zin, dus spontane conversaties zijn echt heel anders dan formele spraak.

Waar wij nu naartoe willen, dat zijn modellen die kunnen verklaren hoe luisteraars met al die eigenschappen van spontane conversaties omgaan, en hoe en waarom sprekers die conversaties produceren zoals ze dat ook doen. Nou, dat is nog heel veel werk, er zijn ook nog heel veel vragen die we moeten beantwoorden, en daarom zal ik hier vast nog wel tot mijn pensioen mee bezig zijn.