

Computer kan beter liplezen dan professional

By **Laurens Verhagen**, www.volkskrant.nl

november 29ste, 2016

Natuurlijke omgeving

De BBC-programma's waren goed voor ruim 118 duizend verschillende zinnen en 17.500 unieke woorden, afkomstig van zo'n duizend sprekers. Ter vergelijking: een ander onderzoeksprogramma in Oxford, genaamd LipNet, moet het doen met duizend zinnen van 34 verschillende sprekers. Dit LipNet gebruikte ook nog eens videomateriaal dat speciaal voor dit doeleinde was opgenomen. Geen presentatoren die door elkaar praten of woorden inslikken, maar mensen die duidelijk articulerende zinnen inspreken.

DeepMind doet zijn werk in een veel natuurlijker omgeving, wat de weg vrijmaakt voor allerlei praktische toepassingen. De onderzoekers noemen in hun artikel de mogelijkheid van het geven van instructies in een mobiele telefoon in een lawaaiige omgeving. Denk aan Siri, maar dan beter. 'Sorry, ik denk niet dat ik u helemaal verstond' is er dan niet meer bij. Een andere toepassing is directe transcriptie van uitgesproken teksten, bijvoorbeeld voor ondertitels. 'Als je dat combineert met spraakherkenning, kom je nog verder dan nu al het geval is', stelt Bohte.

Her en der is al geopperd dat geheime diensten ook wel veel belangstelling zullen hebben voor de liplezende computer. Een vertrouwelijk gesprek in een mensenmassa is er dan niet meer bij. Of is dat sciencefiction? 'Nee hoor', zegt Bohte. 'Dat is prima mogelijk als de kwaliteit van de beelden maar hoog genoeg is.'

