

PDF hosted at the Radboud Repository of the Radboud University Nijmegen

The following full text is a publisher's version.

For additional information about this publication click this link.

<http://hdl.handle.net/2066/42430>

Please be advised that this information was generated on 2019-11-19 and may be subject to change.

Een uitbreiding van het Corpus Gesproken Nederlands

JASMIN-CGN

Automatische spraakherkenning, oftewel herkenning van spraak door computers, wordt onder andere gebruikt in spraakgestuurde toepassingen zoals navigatiesystemen, educatieve computer-programma's, en hulpmiddelen ter ondersteuning van mensen met beperkingen (bijvoorbeeld voor het bedienen van apparatuur met de stem).

| CATIA CUCCHIARINI |

J ongeren
A nderstaligen
S enioren
M achine
I nteractie
N voor het
ederlands

Computers kunnen spraak redelijk goed herkennen als ze “getraind” zijn met grote hoeveelheden spraak, denk bijvoorbeeld aan het informatiesysteem van de NS. Hiervoor zijn spraak-databanken nodig. Voor het Nederlands bestaat bijvoorbeeld het Corpus Gesproken Nederlands (CGN), een verzameling van ongeveer 900 uur gesproken Standaardnederlands, afkomstig van volwassen Nederlandstalige sprekers uit Vlaanderen en Nederland.

Op dit moment kunnen computers de spraak van volwassenen redelijk goed verstaan, maar ze hebben nog steeds veel moeite met spraak van jongeren, anderstaligen en senioren. Dit komt doordat de spraak van deze groepen sprekers heel anders is dan die van volwassenen. Er zijn op dit moment geen databanken met spraak

Een ander bekend probleem bij het inzetten van automatische spraakherkenning om mensen met computers te laten spreken is dat wanneer mensen tegen een computer praten, hun spraak kenmerken vertoont die problematisch zijn voor computers. Bijvoorbeeld aarzelingen, onderbrekingen, versprekingen, schreeuwen en overdreven articuleren.

Beter inzicht in de aard en het voorkomen van deze verschijnselen zou kunnen helpen om te bepalen hoe de technologie aangepast moet worden zodat deze natuurlijke kenmerken van menselijke spraak geen obstakel meer vormen voor automatische spraakherkenning. In het project JASMIN-CGN zullen dan ook opnames gemaakt worden van situaties waarin jongeren, anderstaligen en senioren een dialoog met een computer voeren.

JASMIN-CGN is een samenwerkingsproject tussen de Radboud Universiteit Nijmegen (CLST), de K.U. Leuven (ESAT) en het bedrijf Talking Home.

De resultaten van dit project kunnen gebruikt worden voor onderzoek en voor het ontwikkelen van spraakgestuurde toepassingen voor deze specifieke groepen sprekers.

“Computers hebben nog steeds veel moeite met spraak van jongeren, anderstaligen en senioren.”

van jongeren, anderstaligen en senioren, althans niet voor het Nederlands.

Om aan deze behoefte tegemoet te komen is het project JASMIN-CGN gestart, dat als doel heeft het samenstellen van een databank met spraak van jongeren, anderstaligen en senioren.

CATIA CUCCHIARINI
COÖRDINEERT HET
JASMIN-CGN-PROJECT

Talking HOME

KATHOLIEKE UNIVERSITEIT
LEUVEN

Radboud Universiteit Nijmegen



JASMIN-CGN is een project uit de eerste ronde. Het project loopt tot eind maart 2003. Zie: <http://homes.esat.kuleuven.be/~spch/projects/JASMIN/>.