

## EXPERIMENTEN INZAKE EEN SPRAAKHERKENNINGSSYSTEEM VOOR ESPERANTO

Gedurende enige maanden heb ik met verschillende tests geprobeerd een spraakherkenningssysteem voor Esperanto te realiseren in het kader van het zgn. ALPAZORO-project ('ALPARolu ZOrgoRObotojn'=spreek tegen zorgrobots) waarvan Melanie Maradan en Toon Witkam de initiatiefnemers waren.

(Zie o.a. <http://www.iei.nl/ieieo52.htm> van de website van het Internationaal Esperanto-Instituut: [www.iei.nl](http://www.iei.nl)).

Alle tests en pogingen heb ik gedaan met het computerprogramma *Dragon*, versie 12, van de firma *Nuance*. Dat is een spraakherkenningssysteem waarmee men door een microfoon, die verbonden is met een computer, spreekt en het programma typt dan vrijwel gelijktijdig de tekst op het beeldscherm of op de printer.

Een dergelijk programma is een goed hulpmiddel om na te gaan of gesproken woorden en zinnen door de computer worden 'begrepen', dus waarschijnlijk ook door een robot waarin zich een spraakherkenningssysteem bevindt. (Natuurlijk moet de robot na het 'begrijpen' van de woorden en zinnen vervolgens ook nog handelen: de gegeven opdrachten uitvoeren, maar dat is een stap na de eerste stap: 'begrijpen' (verstaan) van de gesproken woorden en zinnen).

*Dragon* bestaat voor meerdere talen en wordt door vele secretariaten, tijdens conferenties, voor ondertiteling van rechtstreekse/actuele TV-uitzendingen, etc. gebruikt.

In *Dragon* moet men woorden, zinsdelen of zinnen op twee manieren invoeren: de tekst zoals deze moet worden getypt (typvorm) en de tekst volgens uitspraak (uitspraakvorm).

Ik heb twee versies van *Dragon* getest: de Nederlandse en de Italiaanse.

Het idee om de Italiaanse te gebruiken was dat de klanken daarvan erg lijken op de Esperanto-klanken, dus de herkenbaarheid zou erg groot moeten zijn.

In dat geval moet men dus een Esperantowoord (of zinsdeel) invoeren in de typvorm (bijv. *ŝipo, citrono*) en bovendien in de Italiaanse uitspraakvorm, bijv. *scipo, zitrono*.

Helaas heeft het Italiaans voor de Esperantoletters *ĝ, h, j, ĵ, ĥ* en *z* geen letter of lettercombinatie die de esperanto-uitspraak kan weergeven...

De Italiaanse versie van *Dragon* was dus niet geschikt.

(Een andere mogelijkheid zou zijn om een Kroatische, Slowaakse of Tsjechische versie van *Dragon* te gebruiken: die talen zijn tamelijk fonetisch en kennen ook kapjesletters en -klanken. Jammer genoeg heeft *Dragon* geen Kroatische, Slowaakse of Tsjechische versie, slechts een Engelse, Franse, Duitse, Italiaanse, Spaanse en Nederlandse).

Mijn verdere experimenten en pogingen zijn ALLEMAAL met de Nederlandstalige *Dragon*.

### A) Experimenten volgens de zgn. TYKBS-methode

De uitspraakvorm *d* en *t* worden in deze methode *T*; *ĝ* en *ĵ* worden *Y*; *k, g* en *ĥ* worden *K*; *b, v, f* en *p* worden *B*; *c, ĉ, s, ŝ* en *z* worden *S*; *h, l* en *r* worden niets; *ŭ* wordt *w*; *u* wordt *oe*; de *i* wordt *i* of *ie*. (De overige letters veranderen niet).

Dus in dit systeem worden alle klanken die op elkaar lijken één zelfde letter.

Bijv. *ŝipo* (typvorm) wordt *SiBo* (uitspraakvorm).

Bijv. *vazo* (typvorm) wordt *BaSo* (uitspraakvorm).

Bijv. *haroj* (typvorm) wordt *aoj* (uitspraakvorm).

Het is duidelijk dat een dergelijk systeem de kans dat de computer de uitgesproken woorden begrijpt, zeer verhoogt, want als men bijv. de z als s uitspreekt, of de g als k, begrijpt de computer het woord toch.

Maar de uitspraakvorm van *ŝipo* en *cipo* zijn allebei *SiBo*, en de uitspraakvormen van *paŝo* en *vazo* en *bazo* zijn allemaal *Baso*. En de uitspraakvormen van *haloj* en *haroj* zijn beide *aoj*.

In het TYKBS-systeem kan men dus alleen die woorden gebruiken die geen uitspraakvormig equivalent hebben.

Dus in die lexicon mag *SiBo* dus bijv. alleen *ŝipo* en nooit *cipo* betekenen; en *BaSo* bijv. alleen maar *paŝo* en niet *vazo* of *bazo*.

Als men slechts een beperkt lexicon nodig heeft, dus bijv. alleen om opdrachten aan een robot te geven, kan men dus eenvoudigweg uit een reeks gelijkklinkende woorden er één uitkiezen, namelijk het woord waaraan het meeste behoefte is.

### Conclusies:

- De methode functioneert behoorlijk goed.

Meestal mislukken woorden die zeer kort zijn of waarvan de uitspraakvormen volgens de TYKBS-methode zeer sober of afwijkend zijn, bijv. *la* wordt *a*, *pri* wordt *Bie*.

De methode wordt meer succesvol als men in *Dragon* volledige zinnestelsels of zinsdelen invoert, bijv. *mi volas trinki* - *mie Boas TienKie*.

- De gebruikte *Dragon* is de Nederlandstalige versie. Misschien is de TYKBS meer geschikt voor spraakherkenningsystemen die niet gebonden zijn aan een nationale taal.

### B) Experimenten volgens de LETTERGREEP-methode

In deze methode bestaat het *Dragon*-lexicon uitsluitend uit Esperantotalige lettergrepen: typvormig en uitspraakvormig (uitspraakvormig volgens de Nederlandse uitspraak).

Bijv. *ĉir* (typvorm) met *tsjier* (uitspraakvorm) en *kaŭ* (typvorm) met *kaw* (uitspraakvorm).

Bijv. *ĝus* (typvorm) met *dzjoes* (uitspraakvorm) en *ta* (zowel typ- als uitspraakvorm).

Het lexicon bestaat bij deze methode uit (slechts) ca. 1500 lettergrepen.

Een groot voordeel is dat men allerlei en alle woorden kan vormen.

Enigszins hinderlijk is dat men elke lettergreep afzonderlijk en met nadruk moet uitspreken, dus bijv. *promeni* uitspreken als *pro-me-ni*.

Een ander probleem is, dat het niet duidelijk is welke combinatie van lettergrepen een woord is, want alle lettergrepen plakken aan elkaar: *mipromenasenlaparko*.

Dat probleem is gemakkelijk op te lossen: als spatie kan men bijv. een \* (of een ander teken) als typvorm gebruiken en *ieks* (of een ander speciaal woord) als uitspraakvorm. Dus men kan bijv. zeggen: *mie ieks pro me nas ieks en ieks la ieks par ko* en dan zal de volgende getypte tekst verschijnen:

*mi\*promenas\*en\*la\*parko* (Het sterretje kan men later gemakkelijk door een spatie vervangen). Maar een dergelijke manier van spreken is nogal lastig.

### Conclusie

- De methode werkt goed, als men tenminste zeer duidelijk en nadrukkelijk iedere lettergreep van elk woord uitspreekt en tussen de woorden het spatiewoord uitspreekt.

### Experimenten volgens de ALLES-methode

In deze methode bestaat het *Dragon*-lexicon uit alle bestaande Esperantowoorden met alle mogelijke voor- en achtervoegsels, vervoegingen, etc., alle zowel typvormig als uitspraakvormig (uitspraakvormig volgens de Nederlandse uitspraak).

Bijv. manĝi-mandzjie, manĝas-mandzjas, manĝis-mandzjies, manĝante-mandzjante, manĝata-mandzjata, manĝo-mandzjo, manĝon-mandzjon, manĝojn-mandzjojn, manĝaĵo-mandzjazjo, manĝebla-mandzjebla, etc. etc. (Dit systeem wordt overigens ook gebruikt in het Nederlandstalige *Dragon*-lexicon: het complete officiële Nederlandse woordenboek (het '*Groene boekje*') staat er in: eten, eet, at, aten, gegeten, etend, etende, etenswaar, etenswaren, voedsel, maaltijd, maaltijden, etc.).

Bij een dergelijke methode bestaat een tamelijk basaal lexicon uit ten minste 30.000 woorden, elk woord typvormig en uitspraakvormig. Natuurlijk als men de methode alleen maar wil gebruiken voor een beperkt aantal zinnen (bijv. opdrachten voor robots) heeft men veel minder woorden nodig.

### Conclusie:

De methode werkt goed als men tenminste elk woord duidelijk en afzonderlijk uitsprekt, en met een korte pauze tussen elk woord. Als het lexicon minder groot is, zijn de resultaten naar verhouding beter (de foutenkans wordt dan kleiner).

### D) Experimenten volgens de WOORDSTAMMEN-methode

Een speciale eigenschap van Esperanto is dat het uit woordstammen bestaat waaraan men voorvoegsels en achtervoegsels kan toevoegen om woorden te vormen.

In deze methode bestaat het *Dragon*-lexicon uit alle Esperanto-woordstammen en alle mogelijke voor- en achtervoegsels, alle zowel typvormig als uitspraakvormig (de uitspraakvorm volgens de Nederlandse uitspraak).

Voorbeelden: manĝ-mandzj, oj-oj, on-on, ojn-ojn, as-as, is-ies, os-os, aĵ-azj, ant-ant, at-at, ig-ieg, iĝ-iedzj, ebl-ebl, etc.

Op deze manier kan men dus vele woorden vormen, maar voor- en achtervoegsels zijn maar eenmalig nodig in het lexicon.

Eénletterige achtervoegsels zijn echter niet toegestaan. Daarom moeten woorden die zo'n achtervoegsel hebben ook in het lexicon worden opgenomen:

dus: de woordstam *bon* en de woorden *bona*, *bone*, *bono*; de woordstam *danc* en de woorden *danci*, *danco*, *dancu*.

Een dergelijk lexicon zal dus bestaan uit ca. 2.500 woordstammen, ca. 6.000 woorden met -o, etc. en ca. 60 voor- en achtervoegsels.

Aan woordstammen wordt automatisch het volgende achtervoegsel gekoppeld en na een woord volgt automatisch een spatie. Aan een voorvoegsel komt automatisch een woordstam of woord.

Aldus verschijnen spaties automatisch op de juiste manier.

### Conclusie:

De methode werkt zeer goed; hij benut feitelijk de agglutinatiestructuur van Esperanto en is daarom zeer flexibel.

### E) Experimenten volgens de COMPLETE ZINNEN-methode

In deze methode worden complete zinnen in *Dragon* ingevoerd (typevormig en uitspraakvormig, uitspraakvormig volgens de Nederlandse uitspraak).

Bijv.: *la juna knabo manĝas glaciaĵon - la joena knabo mandzjas glatsiazjon*

Deze methode is natuurlijk zeer bruikbaar als men slechts een zeer kleine verzameling zinnen gebruikt, dus bijv. wanneer men een voorbereide lezing houdt en tijdens de lezing de schriftelijke teksten op een scherm projecteert, of om een robot aan te spreken met slechts een gelimiteerd aantal opdrachten, of om standaard-zinnen te genereren in bijv. een akte of contract, etc.

### Conclusie:

De methode werkt erg goed. Deze is slechts bruikbaar voor zeer beperkte spreeksituaties.

### Algemene opmerkingen:

De resultaten en kwaliteit van alle methoden zijn zeer afhankelijk van de kwaliteit van het gebruikte audiosysteem (o.a. microfoon, kwaliteit van de computer (geluidskaart en omvang van het interne geheugen) en vooral van de kwaliteit van de individuele spreker: een duidelijke uitspraak, spraakgeoefendheid, etc. Iedere methode is dus ook nuttig om de eigen uitspraakkwaliteit te onderzoeken en testen, want naarmate die beter is, des te beter zullen de getypte resultaten zijn.

### Algemene conclusie:

Alle genoemde methodes, met uitzondering wellicht van de TYKBS-methode, hebben een bestaand spraakherkenningsysteem nodig, *in casu* Dragon. Alle lettergrepen, woordstammen, woorden of zinnen/zinsdelen moeten in zowel typvorm als uitspraakvorm worden ingebracht. Als men de Nederlandstalige Dragon gebruikt moet de uitspraakvorm Nederlandstalig zijn.

Voor eenvoudige, beperkte taalsituaties is de COMPLETE ZINNEN-methode het beste. Voor gecompliceerde, rijke taalsituaties is de WOORDSTAMMEN-methode het meest geschikt. Als men een methode wil gebruiken om te dicteren (bijv. een brief) dan is de LETTERGREEP-methode zeer bruikbaar.

Het is nuttig en het is de bedoeling deze drie methodes verder te ontwikkelen zodat ook anderen ze kunnen gebruiken. (Lijsten met de WOORDSTAMMEN- en LETTERGREPEN-lexicons zijn bijna compleet en zullen overdraagbaar worden aan nieuwe gebruikers van het systeem).

Naast de ontwikkeling van de methodes met Dragon zou het goed zijn om systemen te onderzoeken die niet van nationale talen afhankelijk zijn en waarvoor de TYKBS-methode mogelijk nuttig en bruikbaar kan zijn. Daarnaast worden er systemen ontwikkeld voor veel andere nationale talen (o.a. het GlobalPhone-project van het Karsruher Institut für Technologie, dat voor die talen databestanden van complete zinnen onderhoudt).

2014, Piet (Petro) Buijnsters