

Samenvatting Scriptie Edwin Wildeboer

Boomstructuren zijn een veelgebruikt middel voor het visualiseren van hiërarchie. Ze worden ook toegepast bij het analyseren van constituenten in zinnen. Normaliter zijn de boomstructuren binair en twee-dimensionaal van vorm en worden over het algemeen gezien als een gemakkelijke representatie van verschillende vormen van hiërarchische informatie. Er doet zich echter een probleem voor als bepaalde informatie als niet-hiërarchisch wordt gezien of betrekking heeft op meerdere niet-onderschikkingen.

Een specifiek toepassingsgebied met hiërarchisch geordende data is de taalkunde, met name de *syntaxis*. In de *syntaxis* worden de relaties tussen de verschillende woordgroepen in een zin bestudeerd en de zinsbouw beschreven.

Dit afstudeeronderzoek sluit aan bij het NWO-project 'The Syntax of Nonsubordination: Parentheses, Appositions and Grafts' (M. de Vries, H. Heringa & M. Kluck) en maakt gebruik van informatievisualisatietechnieken die toegepast worden op syntactische zinsconstructies.

Omdat in de *syntaxis* voor zaken als Coördinatie en Right Node Raising een simpele hiërarchische benadering niet adequaat blijkt te zijn moet daar gezocht worden naar een andere vorm van visualisatie. Daarom is een experiment opgezet om vijf verschillende visualisatievormen van boomdiagrammen (inclusief 3D-bomen) voor het representeren van non-subordinatie en multidominantie te testen. De centrale vraag van het onderzoek is hierbij wat bruikbare alternatieven kunnen zijn voor het visualiseren van deze syntactische boomstructuren.

Twee termen die nader dienen te worden uitgelegd, zijn *non-subordinatie* en *multidominantie*. *Non-subordinatie* is een verzamelterm voor constructies die niet ondergeschikt zijn aan andere constructies. *Multidominantie* heeft betrekking op knopen die verbonden zijn met meerdere, in de bomen hiërarchisch hoger gelegen, knopen. Bij syntactische constructies, waarbij aangenomen kan worden dat er sprake is van non-subordinatie en multidominantie, kunnen zich visualisatieproblemen voordoen. Met name indien we ook hiervoor 2D binaire bomen als representatie willen gebruiken. In de taalkunde zijn er opvattingen die non-subordinatie betrekking laten hebben op syntactische constructies. Hierin komen constituenten voor die niet in een onderschikkende relatie staan tot andere constituenten; dit wordt *nevenschikking* of *coördinatie* genoemd.

De onderzoeksvraag bestaat uit één algemene hoofdvraag die een verdere verdieping bevat in de vorm van enkele deelvragen. Met het experiment wordt gepoogd antwoord te krijgen op deze vragen die hieronder worden genoemd:

Onderzoeksvraag

- Welke vorm van visualisatie is het meest bruikbaar voor de weergave van syntactische constructies waarin non-subordinatie en/of multidominantie voorkomen?

Het voorgaande geeft aanleiding tot het formuleren van enkele vragen die betrekking hebben op de mogelijke bruikbaarheid van de vijf visualisatievormen van syntactische boomstructuren (waarbij sprake is van non-subordinatie en/of multidominantie). Deze vragen worden uitgesplitst naar syntactische categorie (*Coördinatie*, *Right Node Raising* en *Non-constituent Right Node Raising*) en vraagtype (*Constituentie*, *Dominantie*, *Nevenschikking*, *Parallellie* en *Parallele Dominantie*). Ook wordt gekeken of een verschil waarneembaar is tussen de resultaten van de verschillende groepen deelnemers aan het onderzoek, namelijk *leken*, *studenten* en *experts*.

Om antwoord op de bovengenoemde onderzoeksvraag en bijbehorende deelvragen te krijgen, is een empirisch onderzoek in de vorm van een experiment opgezet waaraan 46 deelnemers hebben deelgenomen. Tijdens dit experiment zijn vijf visualisatievormen van syntactische boomstructuren geëvalueerd aan de hand van objectieve en subjectieve beoordelingen. Om de bruikbaarheid van de visualisatievormen te kunnen bepalen, wordt hier getoetst aan de hand van het drietal onderdelen die volgens het ISO bij het toetsen van *usability* (bruikbaarheid) dienen te worden toegepast, namelijk *effectiviteit*, *efficiëntie* en *tevredenheid*. Dit geschiedt met behulp van de te meten waarden die van toepassing zijn bij het toetsen van *usability*. Voor dit onderzoek zijn de drie te meten waarden: foute antwoorden (*error rate*), antwoordtijden en subjectieve beoordeling van de deelnemers.

Voor het experiment zijn drie verschillende groepen (met expert-, basis- en geen kennis op het vlak van taalkunde) benaderd om vragen te beantwoorden met betrekking tot Coördinatie en (Non-constituent) Right Node Raising-constructies en hun voorkeur voor elke visualisatievorm aan te geven.

Uit het onderzoek is gebleken dat de visualisatievorm 2D+ het meest bruikbaar is. In tegenstelling tot de verwachting blijkt uit de resultaten dat voor het representeren van non-subordinatie en/of multidominantie de visualisatievorm 3D+ de minst bruikbare syntactische boomdiagram is.