



Universiteit
Leiden
The Netherlands

Bardaisan's Book of the Laws of the Countries : a computer-assisted linguistic analysis

Bakker, D.

Citation

Bakker, D. (2011, April 20). *Bardaisan's Book of the Laws of the Countries : a computer-assisted linguistic analysis*. Retrieved from <https://hdl.handle.net/1887/17580>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/17580>

Note: To cite this publication please use the final published version (if applicable).

SAMENVATTING

Bardaisans

Boek van de Wetten der Landen:

Een computer-ondersteunde taalkundige analyse.

Het *Boek van de Wetten der Landen* (BWL) is een van de oudste getuigen van de vroege Syrische literatuur. De tekst, een verhandeling in proza gecomponeerd in de 3^e eeuw na Christus, wordt toegeschreven aan de auteur Bardaisan (Edessa, ± 154–222 n. Chr.), alhoewel hij, zo blijkt uit de tekst zelf, in werkelijkheid hoogstwaarschijnlijk op schrift is gesteld door een van zijn leerlingen, Philippus genaamd. Het werk behelst de beschrijving van een theologische discussie tussen Bardaisan en enkele van zijn leerlingen, waarbij verschillende nauw verwante geloofsvraagstukken de revue passeren. Sommige van de onderwerpen zijn van een verrassend actueel gehalte, zoals de vraag of de mens een vrije wil heeft, en of men dient te geloven om te kunnen begrijpen, of juist te begrijpen om te kunnen geloven.

De titel, 'Boek van de Wetten der Landen', dankt het werk aan een omvangrijke passage waarin Bardaisan de verschillende landen in de wereld en hun wetten aandraagt als argument tegen de bewering dat het menselijk handelen onder invloed staat van de sterren, zoals de astrologen beweren. Dat dit niet zo kan zijn, aldus Bardaisan, blijkt uit de observatie dat elk land zijn eigen gewoonten heeft, die ertoe leiden dat alle bewoners zich in hetzelfde land op dezelfde manier gedragen. Dit kan onmogelijk het gevolg van de sterren zijn, want dat zou betekenen dat alle mensen in dat land onder exact dezelfde horoscoop zouden moeten zijn geboren, wat evident niet het geval is. De verklaring van hun gedrag ligt veeleer in hun beschikking over een vrije wil, hun in staat stelt om zichzelf wetten op te leggen. Om zijn argument kracht bij te zetten beschrijft Bardaisan vervolgens een grote hoeveelheid landen en volkeren met hun respectievelijke wetten en gebruiken,

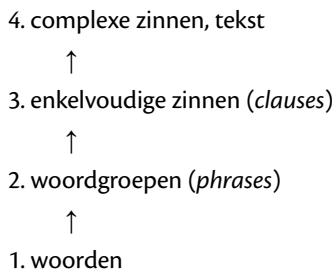
zoals de Brahmanen in India, die geen vlees mogen eten, of de Parthen, die polygaam leven.

De Syrische tekst van BWL is ons slechts in één 6^e- of 7^e-eeuws manuscript overgeleverd, dat halverwege de 19^e eeuw samen met honderden andere handschriften werd ontdekt in het klooster van Deir as-Suryan in de Egyptische Wadi Natrun. Het feit dat er voor die tijd reeds fragmenten uit de tekst bekend waren via Griekse parallellen leidde tot jarenlange discussie over de authenticiteit van het werk, de toeschrijving ervan aan Bardaisan en de taal waarin de tekst oorspronkelijk was geschreven. Pas in de 20^e eeuw werd de consensus bereikt dat die taal het Syrisch was en dat de tekst het gedachtegoed van Bardaisan weerspiegelt. Hiermee vormt het Boek van de Wetten der Landen een werk dat kan worden beschouwd als een representant van 'authentiek' (in plaats van vertaald) Syrisch proza. Wel moet er in dit opzicht rekening worden gehouden met het feit dat het manuscript uit de 6^e of 7^e eeuw stamt, dus grofweg drie eeuwen na de compositie van de oorspronkelijke tekst. We hebben te maken met een kopie (die zelf ook weer kan teruggaan op een kopie, etc.), die derhalve zowel wat taalgebruik als wat inhoud betreft kan afwijken van het origineel.

Het doel van deze dissertatie was het onderwerpen van de Syrische tekst van BWL aan een gedetailleerde taalkundige analyse. Voor deze analyse heb ik gebruik gemaakt van computer-ondersteunde methoden en modellen die zijn ontwikkeld in de projecten CALAP en *Turgama* in samenwerkingsverband tussen de Werkgroep Informatica Vrije Universiteit en het Peshitta Instituut Leiden. Die methoden en modellen zijn speciaal toegespitst op onderzoek naar oude talen, zoals het Hebreeuws en Syrisch, die als inherente problematiek hebben dat zij niet langer gesproken worden en dus geen *native speakers* meer hebben. Aangezien het onderzoek naar dergelijke dode talen noodzakelijkerwijs is aangewezen op overgeleverde, geschreven teksten, hanteert het CALAP/*Turgama*-model een corpuslinguïstische methode, die zich zo veel mogelijk richt op distributionele verschijnselen die in de tekst waarneembaar zijn, met zo min mogelijk interferentie van externe kennis over de taal. Dat dit ideaal van een zuiver op de data gebaseerde analyse in de praktijk slechts kan worden benaderd, blijkt uit het feit dat het CALAP/*Turgama*-model enkele fasen bevat waar menselijke inmenging vereist is: in het opstellen van een set morfologische basisregels in de vorm van

een *word grammar*, alsook in een interactief analyseproces waar mens en computer elkaar wederzijds aanvullen (Hoofdstuk 1, § 2.2).

De computerprogramma's die voor de analyse worden gebruikt stellen de onderzoeker in staat om een tekst nauwgezet te ontleden in taalkundige eenheden: morfemen, woorden, zinsdelen, etc. Op grond van het soort en de omvang van de taalkundige eenheid waarnaar gekeken wordt kent het analyseproces vier afzonderlijke stadia, die corresponderen met vier taalkundige niveaus: woordniveau, woordgroepniveau, zinsniveau en tekstniveau. Deze niveaus worden achtereenvolgens geanalyseerd volgens een *bottom-up* benadering, waarbij de data verkregen op een lager niveau telkens wordt geïntegreerd in het niveau erboven. Op het eerste, laagste niveau wordt gekeken naar de opbouw van woorden uit morfemen (lexemen en hun inflectie); op het niveau daarboven naar de opbouw van woorden in woordgroepen (*phrases*); vervolgens naar de opbouw van woordgroepen in enkelvoudige zinnen (*clauses*), en tenslotte naar de opbouw van enkelvoudige zinnen in complexe zinnen en grotere tekstpassages:



De vier taalniveaus, van beneden naar boven

De analyse van deze vier niveaus verloopt halfautomatisch, in een interactief proces tussen gebruiker en computer: op basis van eerder gemaakte keuzes door de gebruiker doen de computerprogramma's suggesties voor de analyse van een vorm of constructie, die de gebruiker op grond van zijn taalkundig inzicht accepteert, weigert of aanpast, waarna die beslissingen weer worden verwerkt in de lijst van suggesties die de computer kan doen. Na voltooiing van de analyse van een bepaald taalniveau wordt de verkregen data vervolgens gebruikt als input voor de analyse van het niveau erboven. Deze gecontroleerde wisselwerking tussen computer en gebruiker garandeert

niet alleen een hoge mate van consistentie, maar biedt tevens een effectieve mogelijkheid tot zelfcontrole: zodra een keuze van de onderzoeker op protest van de computer stuit kan dit duiden op een probleem in de programmatuur, maar ook op een taalkundige eigenaardigheid die nadere studie verdient. Aldus 'leren' de mens en de computer van elkaar.

Wanneer de onderzoeker op deze manier zijn hele corpus doorlopen heeft is het resultaat een zeer uitgebreide database van taalkundige informatie over de desbetreffende tekst. Die database dient vervolgens als basis voor het daadwerkelijke taalonderzoek, waarbij de onderzoeker de data met behulp van geavanceerde zoekopdrachten kan inventariseren, ordenen en bestuderen.

In mijn dissertatie heb ik, met deze database als uitgangspunt, de eerste drie taalniveaus (woordniveau, woordgroepniveau en zinsniveau) van BWL geanalyseerd, waarbij ik me per niveau telkens op een selectie uit de vele taalkundige fenomenen heb gericht.

Uit het onderzoek is gebleken dat wat voor de menselijke input op taalkundig niveau geldt, des te meer van toepassing is waar het kennis van de tekst zelf betreft. Zelfs met behulp van geavanceerde rekenmethoden is een adequate analyse van een corpus niet mogelijk op grond van de grammaticale taalverschijnselen alleen. Bij onduidelijke of dubbelzinnige vormen moet de onderzoeker, wanneer de grammaticale gegevens niet toereikend zijn, voor zijn oordeel afgaan op context en inhoud van de tekst of op zijn algemene kennis van de werkelijkheid. Met andere woorden: bij de taalkundige analyse van een corpus is er voortdurend sprake van wisselwerking tussen taalkunde en filologie.

De relevantie van deze interactie voor de analyse van Bardaisans BWL blijkt niet alleen uit het feit dat het onderwerp van de tekst de linguïstische kenmerken ervan beïnvloedt (bijv. de veelheid aan eigen-, volks- en geografische namen), maar ook in diverse gevallen van noodzakelijke filologische overwegingen voor een vellen van een adequaat taalkundig oordeel, bijvoorbeeld in de interpretatie van de *waw* in ܡܝܢ ܕܥܝܪܐ als *waw explicativum* (in Syrië, **namelijk** in Edessa: Hfst. 2, n. 22), de beslissing met betrekking tot de identiteit van de interjectie ܐܝܬܐ (Hfst. 3, n. 164), de analyse van de constructie ܐܝܬܐ ܕܡܝܢ ܕܥܝܪܐ (Hfst. 3, § 2.1.2.4) of het bepalen van thema en rhema in een zin (Hfst. 4).

Ofschoon de onderhavige studie zich op de taalkundige aspecten van BWL richtte, was dus een beknopte beschrijving van de inhoud van de tekst noodzakelijk (Hfst. 1, § 1), alsook een overzicht van eerder onderzoek inzake de herkomst en authenticiteit ervan. Hoofdstuk 1 bevat tevens een uiteenzetting van de problemen bij het vaststellen van de textuele basis, wanneer—zoals in dit geval—slechts één tekstgetuige beschikbaar is. Vastgesteld diende te worden of er sprake is van verschillende tekstlagen en hoe deze zich chronologisch tot elkaar verhouden.

In Hoofdstuk 2 wordt uitgelegd dat de eigenaardigheden van BWL op het gebied van spelling en vormleer kenmerkend zijn voor het Syrische taaleigen van de 6^e of 7^e eeuw. Deze komen dus overeen met de tijd waarin het handschrift is gekopieerd en niet zozeer met de tijd waarin de tekst oorspronkelijk geschreven werd, iets waar terdege rekening mee moet worden gehouden bij verder taalhistorisch onderzoek.

Voorts is uit mijn onderzoek van de spelling gebleken dat deze kenmerken door drie verschillende factoren zijn beïnvloed, waarbij telkens twee principes een rol spleen: 1. encyclopedie en taalkunde (het feit dat één woord verschillende semantische referenten kan hebben; bijv. de naam ܐܒܓܪ *Abgar*, die naar verschillende personen kan verwijzen; of dat dezelfde referent door verschillende termen kan worden aangeduid; bijv. ܐܠܗܐ en ܐܠܡܐ voor *God*, of ܐܡܝܢ en ܐܝܢܐ voor *India*); 2. orthografie en morfologie (bijv. de vraag of twee varianten van hetzelfde woord, zoals ܠܬܪܐ en ܠܬܪܐܐ *Hatra*, slecht een verschil in spelling weergeven of duiden op het bestaan van twee verschillende lexemen met verschillende uitspraak); en 3. zelfstandige naamwoorden en eigennamen (het feit dat bijv. de namen voor de dierenriemtekens in oorsprong gewone zelfstandige naamwoorden zijn—ܠܬܪܐ *krab*, ܠܬܪܐ *lam*, enz.—en pas in de context tot eigennamen worden (*Kreeft*, *Ram*, enz.)).

Hoofdstuk 3 behelst een gedetailleerd overzicht van alle woordgroeptypen en -patronen die in BWL voorkomen. Daarbij heb ik niet alleen een lijst opgesteld van de afzonderlijke elementen die een woordgroep kunnen vormen, maar heb ik tevens de mogelijke manieren onderzocht waarop dergelijke elementen (ingedeeld naar 'hoofd' en 'extensie') worden gebruikt om zinsdelen te vormen. Dit overzicht bleek zeer inzichtelijk in het licht van de dichotomie tussen taalkundige theorie en waargenomen praktijk. Waar de

recursiviteit ('herhaalbaarheid') van talige structuren in theorie eindeloze combinaties zou kunnen opleveren, blijkt er in de praktijk steeds een bovenglimiet te bestaan voor de complexiteit van de constructies die in een corpus worden aangetroffen. Zo is gebleken dat in BWL een nominale woordgroep (*noun phrase*, NP) ten hoogste zes parallelle hoofden kan hebben, dat een dergelijk hoofd uit niet meer dan twee opeenvolgende nomina in de status constructus kan bevatten, terwijl het aantal afzonderlijke extensies dat een NP kan hebben maximaal vier is. De mate waarin de extensie van woordgroepen recursief is (d.w.z. het aantal malen dat een extensie zelf weer een extensie heeft) blijkt ten hoogste drie te zijn: een extensie van een extensie van een extensie van een NP-hoofd.

Tenslotte heb ik in Hoofdstuk 4 onderzocht hoe zinsdelen in BWL zich binnen een zin gedragen. Daarbij heb ik vooral gelet op het verschijnsel dat een zinsdeel vooraan in een werkwoordelijke zin wordt geplaatst (*fronting*) en de vraag gesteld wat de mogelijke pragmatische implicaties daarvan zijn (bijv. het markeren van thema en rhema) en het effect op *participant activation* (introducering, continuering of reactivering). Op grond van de vaststelling dat de taal van BWL tamelijk strak vasthoudt aan de standaard volgorde VSO (*verb - subject - object*) in het Semitisch, heb ik getracht afwijken van deze volgorde te verklaren bij vooropplaatsing (d.w.z. plaatsing vóór het werkwoord) van één of meer zinsdelen. Het is gebleken dat er in de taal van BWL verschillende correlaties bestaan tussen de grammaticale functie van vooropgeplaatste zinsdelen en hun pragmatische functie. Aange-toond wordt bijvoorbeeld dat een vooropgeplaatst subject een significante neiging heeft als thema gemarkeerd te worden, terwijl vooropgeplaatste objecten juist veelal het rhema markeren. In tegenstelling tot andere zinsdeeltypen worden adjuncten (in het bijzonder temporele) vaak in pre-verbale positie geplaatst, een verschijnsel dat goed lijkt aan te sluiten bij hun inherente functie van het beschrijven van een toestand, wat in de Functionele Grammatica *setting* wordt genoemd. Nog duidelijker dan hun functie als markering van thema en rhema is de rol die vooropgeplaatste zinsdelen hebben ten behoeve van *participant activation*. Dit blijkt vooral uit de relatie tussen de 'zwaarte' van het type woordgroep waaruit het vooropgeplaatste zinsdeel bestaat, en de 'sterkte' van het soort *participant activation*: in het algemeen zullen vooropgeplaatste zinsdelen die uit NP's bestaan, de sterkste activering

bewerkstelligen, namelijk introducering. Demonstratieve pronominale woordgroepen behelzen een zwakkere activering, namelijk reactivering, terwijl persoonlijke voornaamwoordgroepen veelal de zwakste activering opleveren, namelijk continuering. Voorts heb ik aangetoond dat indien meer dan één constituent vooropgeplaatst is, deze zonder uitzondering geplaatst worden in de relatieve volgorde:

subject – complement – object – adjunct / modifier.

In mijn onderzoek naar het taaleigen van BWL heb ik uit een veelheid van mogelijke verschijnselen een keuze moeten maken. Daarbij is de aandacht vooral uitgegaan naar verschijnselen die tot dusverre slechts oppervlakkig bestudeerd waren, zoals de interne structuur van zinsdelen en de kenmerken van de verbale zin, die in vergelijking met de grondig bestudeerde non-verbale zin tamelijk onderbelicht is gebleven.

Verder onderzoek naar de taal van BWL is zeker mogelijk en gewenst om deze unieke getuige van de vroege Syrische literatuur recht te doen. Zo blijft onder andere een taalkundig verantwoorde en actuele wetenschappelijke editie van de tekst een desideratum. Aan dit doel hoop ik met dit proefschrift een bijdrage te hebben geleverd en een beter inzicht te hebben gegeven in zowel diverse taalkundige kenmerken van BWL als in de taal van de vroege Syrische literatuur in het algemeen.

