

De documentatie en beschrijving van bedreigde talen en hun bijdrage aan het onderzoek naar taalkundige universalia.

taalbeschrijving, taaltypologie, documentaire taalkunde, Myenetalen, Bantoetalen

description linguistique, typologie linguistique, linguistique documentaire, langues myènè, langues bantoues

language description, linguistic typology, documentary linguistics, Myene languages, Bantu languages

samenvatting

Deze bijdrage illustreert een aantal recente evoluties in de taaltypologie en de documentaire taalkunde aan de hand van een lopend onderzoeksproject over de Gabonese Bantoetaal Myene. Ze bespreekt de noodzaak van een meer respectvolle omgang met primaire gegevens en geeft aan hoe dit kan. Een korte bespreking van het grammaticale gedrag van relatiefzinnen in het Myene toont aan waarom we meer, betere en beter gedocumenteerde beschrijvingen van de talen van de wereld nodig hebben.

résumé

Dans la présente contribution, un projet de recherches sur la langue bantoue gabonaise myènè sert d'illustration d'un nombre d'évolutions récentes en linguistique documentaire et en typologie linguistique. La récolte, la sauvegarde et la mise à disposition des données primaires sur les langues en voie de disparition méritent plus d'attention dans les recherches de terrain. Une discussion concise du comportement grammatical des phrases relatives en myènè démontre l'intérêt d'un effort de documentation et de description des langues du monde pour la linguistique générale.

abstract

This paper uses an ongoing documentation project on the Myene language of Gabon to illustrate some recent evolutions in linguistic typology and documentary linguistics. It discusses the need for a more respectful approach to primary data and provides a brief overview of best practices for doing so. It also points out why we need more, better and more correctly documented descriptions of the languages of the world.

1. Inleiding

Deze bijdrage illustreert een aantal recente evoluties in de documentaire taalkunde en de taaltypologie aan de hand van het lopende project *Comparative Documentation of the Myene Language Cluster*. Dit driejarige onderzoeksproject wordt gefinancierd door de *Endangered Language Documentation Programme*, zelf een onderdeel van het *Hans Rausing Endangered Language Project* (HRELP, www.hrelp.org), verbonden aan de *School of Oriental and African Studies* in Londen. Samen met het Duitse programma *DoBeS* (Dokumentation Bedrohter Sprachen), is HRELP een concreet antwoord op de vaststelling dat een aanzienlijk

percentage van de talen van de wereld dreigt te verdwijnen zonder enig spoor na te laten. Een pessimistische schatting van UNESCO heeft het over vijftig à negentig procent van de vandaag gesproken talen tegen het einde van de eenentwintigste eeuw. Het engagement van HRELP en DoBeS bestaat erin de documentatie van een groot aantal bedreigde talen te financieren.

In Sectie 2 situeer ik het Myene geografisch en genealogisch, beschrijf ik in welke mate het met uitsterven bedreigd is en leg ik kort uit hoe het zover is kunnen komen. Sectie 3 introduceert het doel en de methodes van de relatief nieuwe discipline van de documentaire taalkunde. Sectie 4, tot slot, beargumenteert kort de relevantie van de beschrijving en documentering van de taalkundige diversiteit voor de algemene taalkunde en bespreekt, ter illustratie, enkele voorlopige onderzoeksresultaten van het Myeneproject.

2. Myene

Myene is de verzamelnaam voor zes dialecten die traditioneel gesproken worden in een aantal regio's in het westen van Gabon, nl. het Adjoemba, het Enenga, het Galwa, het Mpongwe, het Nkomi en het Oroengoe. Merk op dat de term *dialect* hier gebruikt wordt in de strikte, taalkundige zin van het woord, nl. een geografisch gedefinieerde variant van een taal, waarbij *taal* op zijn beurt (niet geheel onproblematisch) gedefinieerd kan worden als een verzameling onderling verstaanbare dialecten. Het Mpongwe wordt of werd gesproken in en rond de hoofdstad Libreville, het Oroengoe in de regio van Port-Gentil, het Nkomi rond de lagune van Fernan Vaz en de andere drie dialecten in de regio van het stadje Lambarene, dat onder andere bekend is als de plaats waar Albert Schweizer zijn hospitaal stichtte.

Het Myene behoort tot de familie van de Bantoetalen, die zo groot is dat aan individuele talen doorgaans niet alleen met hun naam maar ook met een zogenaamd Guthrienummer gerefereerd wordt, een combinatie van een hoofdletter en twee (soms drie) cijfers. De letter staat voor één van de zestien geografische zones waarin het Bantoesprekende gebied onderverdeeld is. Het eerste cijfer slaat op een subgroep en het tweede cijfer op een individuele taal. Het Myene draagt het nummer B11 in deze referentiële classificatie. Er zijn geen andere talen in de B10 groep. Zone B vormt samen met zone A en delen van zones C en H de noordwestelijke tak van de Bantoefamilie, de tak die het dichtst bij het oorsprongsgebied van de Bantoetalen gesitueerd is (ergens op de grens tussen het huidige Kameroen en Nigeria) en waar de taalkundige fragmentatie en de typologische variatie het grootst zijn.

Van tientallen (misschien honderden) Afrikaanse talen weten we alleen dat ze bestaan, of beschikken we hoogstens over een al dan niet betrouwbaar lijstje basiswoordenschat. Onze kennis van het Myene was bij de aanvang van het documentatieproject heel wat beter. We beschikken onder andere over een uitgebreid woordenboek van het Mpongwe, door de Gabonese geestelijke André Raponda Walker (Raponda Walker 1934), maar dit naslagwerk heeft de afwezigheid van toonmarkering als groot nadeel. Daarnaast zijn er verschillende kleinere publicaties over aspecten van vooral de historische grammatica van Myenevarianten (bijvoorbeeld Grégoire & Rekanga 1994, Mouguiama-Daouda 1990), en recent ook een beschrijvende grammatica van het Oroengoe door Odette Ambouroue, die moedertaalspreker is van deze variant en als postdoctoraal onderzoeker deelneemt aan het huidige documentatieproject (Ambouroue 2007). Tot slot zijn er een aantal studies van het tonale systeem van het Myene (deel 2 en 3 van Ambouroue 2007, Grégoire 1991 en

Philippon & Puech). Deze studies van de tonaliteit zijn erg uiteenlopend in aanpak en geen enkele bereikt bevredigende conclusies, wat te wijten is aan de buitengewone complexiteit van het Myene toonsysteem. Grégoire en Ambouroué maken gebruik van een klassieke aanpak met onderliggende tonen waarvan de realisatie bepaald wordt door een opeenvolging van geconditioneerde toonregels, terwijl Philippon en Puech werken met beperkingen op mogelijke of optimale toonschema's binnen bepaalde prosodische domeinen. Het onderscheid tussen hoog en laag is equipollent bij Grégoire en Ambouroué (d.w.z. dat een structurele hoge toon contrasteert met een structurele lage toon) en privaat bij Philippon en Puech (een structurele hoge toon contrasteert met de afwezigheid van een toon, er zijn dus geen structurele lage tonen).

Op basis van mondelinge verslagen waren we er bij de aanvraag van het documentatieproject vanuit gegaan dat het Adjoemba alleen oudere sprekers had, dat het Enenga uitgestorven was en dat de andere dialecten niet bedreigd waren. Onze bedoeling was dus zoveel mogelijk materiaal over het Adjoemba te verzamelen, en de andere dialecten (behalve het Enenga) volledig te documenteren, met inbegrip van de taalverwerving. Voor dit laatste zouden we een aantal babies en peuters drie jaar lang op geregelde tijdstippen filmen en opnemen, om de transcripties achteraf met een specialist taalverwerving te analyseren. Na een eerste onderzoeksverblijf in Gabon bleken de verslagen over de situatie van het Adjoemba en het Enenga correct te zijn, hoewel we na lang zoeken nog één spreker van het Enenga gevonden hebben, een vrouw van rond de tachtig. De andere dialecten blijken echter ook alle met uitsterven bedreigd te zijn. Na twee jaar hebben we bijvoorbeeld nog geen enkel kind kunnen vinden dat geschikt is om deel te nemen aan een studie over de verwerving van het Myene. In alle plaatsen die we voorlopig bezocht hebben, is de moedertaal van kinderen wier ouders Myenetaalig zijn vandaag de dag immers het Frans.

Sinds de koloniale periode worden het Myene en de andere oorspronkelijke Gabonese talen systematisch uit het onderwijs, de media en de administratie geweerd, ten voordele van het Frans. In het onderwijs gebeurde dit tot voor kort (nu hoeft het niet meer) onder andere door middel van het systeem van het symbool. Wie betrapt wordt op het spreken van het Myene tijdens de lessen of daarbuiten, moet een "symbool" dragen. Een leerling kan zich alleen van het symbool bevrijden door een andere leerling te betrappen en te verklikken. Wie het symbool op het einde van de dag draagt, krijgt een straf. Dit perverse systeem vernietigt de loyaliteit tussen leden van een taalgemeenschap en verbant het gebruik van een "minderheidstaal" naar de familiale sfeer, maar het bedreigt op zich niet reeds het voortbestaan van een taal. Kinderen verwerven hun moedertaal immers in de eerste levensjaren, vooraleer ze de schoolgaande leeftijd bereikt hebben. Het voortbestaan van een taal is bedreigd als ouders hun kinderen niet langer in hun moedertaal opvoeden, maar in een andere, dominante taal. Een typische combinatie van redenen zet Myenesprekende ouders ertoe aan hun kinderen van bij de geboorte in het Frans op te voeden. Het Myene wordt al dan niet bewust en al dan niet terecht beschouwd als een obstakel bij de ontwikkeling van het kind tot een succesvol lid van de moderne maatschappij. De dominante taal, het Frans, is overal aanwezig en beschikbaar, zodat het bijvoorbeeld het natuurlijke communicatiemiddel wordt in gemengde huwelijken en ook in de steden, waar sprekers van verschillende Gabonese talen samenwonen. Radio en televisie zorgen er bovendien voor dat het Frans zelfs in eentalige rurale gebieden alomtegenwoordig is in de familiale sfeer. Als het Myene sneller uitsterft dan

sommige andere Gabonese talen, dan is dat wellicht het gevolg van het feit dat de Myenebevolking sterk verstedelijkt is. Uit onze ervaring blijkt dat ouders zich niet helemaal bewust zijn van de gevolgen van hun “keuze” voor het voortbestaan van het Myene. Als we hen vragen of hun kinderen Myene spreken, is het paradoxale antwoord steevast “jammer genoeg niet.” De teloorgang van het Myene wordt ervaren als een pijnlijk, maar onafwendbaar verlies en onze inspanningen om de taal te documenteren kan bij alle generaties op steun en sympathie rekenen.

3. Het hoe en waarom van de documentaire taalkunde

Taalkundigen die de documentaire taalkunde als een aparte discipline erkennen, situeren haar origine vaak in het artikel *Documentary and descriptive linguistics* van Nikolaus Himmelmann (1998). Himmelmann pleit hierin voor een striktere scheiding tussen de twee basisactiviteiten van de beschrijvende taalkunde: enerzijds het verzamelen van primaire gegevens en anderzijds hun grammaticale analyse. De eerste activiteit heeft volgens Himmelmann nood aan een meer solide theoretische en methodologisch basis, waarvoor hij zelf de aanzet geeft. In deze sectie overlopen we een aantal redenen om zorgvuldiger met onze primaire bronnen om te springen dan nu doorgaans het geval is. Daarna bespreek ik enkele zogenaamde *best practices* die in het kader van de documentaire taalkunde ontwikkeld zijn en die we toepassen in het Myeneproject.

De grammaticale beschrijving van talen zonder geschreven traditie is gebaseerd op twee types primaire bronnen: uitgelokt taalgebruik (o.a. door middel van lexicale en grammaticale enquêtes) en spontaan taalgebruik. Het eerste type, ook elicitatie genoemd, is vooral bij het begin van een descriptief project belangrijk, en later om zeldzame fenomenen te exploreren en gaten op te vullen. Het tweede type is uiteraard minder gecontroleerd en moeilijker te verwerven en te beheren, maar het is ook veel rijker en in veel gevallen betrouwbaarder. Onderzoek op basis van corpora van spontaan taalgebruik is vaak noodzakelijk om de taakverdeling tussen grammaticale constructies te begrijpen. De meeste beschrijvende linguïsten zien het als hun voornaamste taak om de massa primaire gegevens te reduceren tot een grammaticaal systeem. Doorgaans is het het resultaat van dit reductieproces dat gepubliceerd wordt, terwijl de primaire gegevens in het beste geval in een persoonlijk of lokaal toegankelijk archief belanden. Als gebruiker van taalbeschrijvingen weten we daarom zelden op hoeveel gegevens een beschrijving gebaseerd is, of deze gegevens spontaan geproduceerd zijn of ontlokt en wat de kwaliteit is van hun interpretatie. We vernemen bovendien alleen iets over die aspecten die de linguïst relevant achtte op het moment van de beschrijving. Met andere woorden, als we als veldwerkers laks omspringen met onze primaire gegevens en alleen de resultaten van onze analyses beschikbaar maken, dan is ons onderzoek niet of nauwelijks reproduceerbaar en dus niet falsifieerbaar. Hiaten in onze analyse kunnen later niet opgevuld worden door onderzoekers met een ruimer of ander interesseveld en/of betere methodes en technieken, tenzij ze zelf opnieuw van nul beginnen met het verzamelen van primaire gegevens. Dit laatste is niet alleen duur en tijdrovend, in de huidige context van massale taalsterfte zal het vaak onmogelijk zijn. Alles wat we nu niet archiveren, dreigt dus voorgoed verloren te gaan. Een laatste, maar niet minder belangrijke reden om primaire gegevens over talen zonder geschreven traditie zorgvuldiger te verzamelen, bewaren en

ontsluiten is dat ze voor de taalgemeenschap zelf veel relevanter zijn dan de beschrijvende grammatica's die erop gebaseerd zijn.

Als we als veldwerkers eenmaal overtuigd zijn van de noodzaak om primaire gegevens te verzamelen die onze onmiddellijke interesses overstijgen en ze bovendien nu en in de toekomst beschikbaar te maken voor collega's en leden van de taalgemeenschap, dan rest de vraag hoe we dat het best aanpakken. Vooruitgang in de documentaire taalkunde heeft geleid tot een consensus rond een hele reeks aanbevelingen om de volledigheid, de duurzaamheid en de toegankelijkheid van primaire taalkundige gegevens te verzekeren.

In het Myeneproject proberen we op twee vlakken volledigheid te bereiken. Enerzijds door een maximale diversiteit na te streven in de types taaluitingen die we documenteren, en anderzijds door de informatie die we verzamelen zo rijk mogelijk te maken. Wat het eerste punt betreft, zorgen we ervoor het taalgebruik van verschillende soorten sprekers te documenteren: jongeren versus ouderen, plattelandsbewoners versus stedelingen en sprekers van verschillende dialecten. Daarnaast proberen we een zo volledig mogelijke documentatie van discourstypes en communicatieve situaties te verzamelen. Om diversiteit op dit vlak te beoordelen, maken we gebruik van een schaal die communicatieve situaties rangschikt van minder gepland (bv een uitroep van verbazing) naar meer gepland (bv. een zorgvuldig voorbereide toespraak). De minst geplande taaluitingen zijn het moeilijkst te documenteren. Andere parameters van variatie zijn het aantal sprekers in een situatie (we nemen monologen, dialogen en groeps gesprekken op), het teksttype (literair, conversationeel, procedureel) en het onderwerp. Wat het laatste betreft, doen we een extra inspanning om bedreigde kennis te documenteren, omtrent onder andere technologieën, de natuurlijke omgeving, maatschappij en geloof, waarbij we erop toezien het project niet in folkloristisch vaarwater terecht te laten komen. Bij het maken van geluidsopnames, streven we de best mogelijke kwaliteit na, zodat alle opnames in de toekomst gebruikt kunnen worden voor akoestisch-fonetische analyses. Video-opnames zorgen voor de toevoeging van contextuele informatie die relevant kan blijken in toekomstig onderzoek, bijvoorbeeld naar de relatie tussen taal, mimiek en gebaren.

De kwaliteit van opnames is eveneens relevant voor de duurzaamheid van documentatie, wellicht het aspect van de verzameling van primaire gegevens waarover het minst nagedacht is voor de opgang van de documentaire taalkunde. Technologieën voor de opname van geluid volgen elkaar in snel tempo op (magneetband, DAT, minidisc, solid state, ...). Het gevolg hiervan is dat bestaande opnames relatief snel ontoegankelijk worden. Opnames op magneetband hebben het voordeel dat ze potentieel van uitstekende kwaliteit zijn (analoge opname), maar magneetband is een kwetsbare drager en toestellen waarmee ze gelezen kunnen worden, zijn steeds moeilijker te verkrijgen. De levensduur van sommige nieuwere technologieën, zoals minidisc, is nog kortstondiger. Het wordt snel een probleem om voor elk type drager in een geluidsarchief een bijhorende lezer te hebben. Bij digitale opname-apparatuur vormt software een bijkomend probleem. Propriëtaire bestandstypes (audio of video) kunnen vaak alleen gelezen worden met apparatuur van hetzelfde merk als dat waarmee ze opgenomen werden. Bovendien zijn veel digitale bestandstypes niet geschikt voor fonetisch onderzoek, omdat ze gecompriemd zijn, zodat potentieel relevante informatie verloren gaat. Op dit vlak is de taak van de documentaire taalkunde er vooral een van bewustmaking. De oplossing van deze problemen ligt immers voor de hand. Het volstaat om bij het maken, digitaliseren of opslaan van audiobestanden geen gebruik te maken van

gecomprimeerde of propriëtaire bestandstypen. In het Myeneproject gebruiken we het .wav-formaat. Dezelfde waakzaamheid is geboden bij alle andere software die gebruikt wordt. Tekstbestanden die in een onbekend en/of obsoleet lettertype opgesteld zijn, zijn voor altijd onleesbaar, tenzij ze de UNICODE standaard respecteren. Er zijn steeds meer Unicode-lettertypes beschikbaar, zodat het ook hier volstaat om bij het begin van een onderzoeksproject een paar juiste keuzes te maken. Producenten van populaire tekstverwerkers (zoals Word) of databaseprogramma's (zoals Filemaker) garanderen niet dat bestanden leesbaar zullen zijn in nieuwere versies van het programma. Ook hier is het dus aan te raden de voorkeur te geven aan niet propriëtaire software.

Het probleem van de duurzaamheid van documentatie is eigenlijk een onderdeel van het ruimere probleem van de toegankelijkheid. Primaire gegevens verzameld tijdens een onderzoeksproject die de duur van dat project overleven, hebben alleen zin als ze toegankelijk en interpreteerbaar zijn. Omdat taal een symbolisch systeem is, zijn opnames van talen zonder geschreven traditie alleen bruikbaar als ze vergezeld zijn van een vertaling en een gedetailleerde annotatie. Het is daarbij belangrijk dat vertaling en annotatie via tijdscode verbonden zijn aan de bijhorende opname. Om dit te bereiken, gebruiken we in het Myeneproject het annotatieprogramma ELAN, ontwikkeld door het Max Planckinstituut voor psycholinguïstiek in Nijmegen. Het is bovendien van het grootste belang dat een documentatie vergezeld gaat van een bestand met metagegevens. In het Myeneproject bestaan de metagegevens enerzijds uit een tekstbestand met uitleg over de structuur van de documentatie, de opbouw van annotatiebestanden en de gebruikte software en anderzijds uit twee uitgebreide gegevensbestanden. Een daarvan bevat gedetailleerde biografische informatie over de sprekers die in de opnames voorkomen. De andere bevat alle potentieel relevante gegevens over opnamesessies, zoals plaats en tijdstip, deelnemers (met een verwijzing naar het sprekersbestand), onderwerpen en discourstypes, instellingen van de opname-apparatuur, eventuele moeilijkheden bij het opnemen of andere relevante contextuele informatie, een samenvatting van de inhoud, een overzicht van de gerepresenteerde discourstypes, en vooral een lijst van alle bestanden die tijdens een sessie aangemaakt zijn (beeld, audio, video) of ermee verbonden zijn (annotatie, vertaling). Deze bestanden hebben overigens alle een mnemotechnische naam, wat de metagegevens extra robuust maakt. De naam van elk bestand begint met drie letters die naar het dialect verwijzen (bijvoorbeeld *ENE* voor het Enenga), gevolgd door de datum van de opname (JJ-MM-DD), de initialen van de onderzoeker die het bestand gemaakt heeft, een volgnummer en de extensie van het bestandstype.

De laatste schakel in de creatie van een duurzame en toegankelijke documentatie is de archivering. Op het einde van het Myeneproject zal de verzamelde documentatie gearchiveerd worden in het in Londen gevestigde ELAR-archief. ELAR (Endangered Languages Archive) belooft een veilige archivering op lange termijn en verbindt zich ertoe de verzamelde documentatie ter beschikking te stellen van alle geïnteresseerden onder de voorwaarden gesteld door de gedocumenteerde sprekers en de makers van de taaldocumentatie. Deze voorwaarden kunnen er bijvoorbeeld in bestaan dat bepaalde opnames pas na het overlijden van de spreker openbaar gemaakt wordt, of dat ze alleen voor onderzoekers toegankelijk zijn, of, integendeel, alleen voor leden van de gedocumenteerde taalgemeenschap. De controle van leden van de gedocumenteerde taalgemeenschap over de toegankelijkheid van de

documentatie is overigens één van de vele ethische aspecten van taaldocumentatie waarvoor binnen de documentaire taalkunde strikte richtlijnen geformuleerd zijn.

4. De essentiële bijdrage van taaldocumentatie aan de ontwikkeling van de theoretische taalkunde: een voorlopig resultaat van het Myeneproject

Er zijn veel redenen om bedreigde talen grondig te documenteren. Een goede documentatie kan bijvoorbeeld het uitsterven van een taal afwenden, als de taalgemeenschap dat wenst, omdat ze de basis kan vormen voor een revitaliseringsprogramma. Mijn persoonlijke motivatie is vooral wetenschappelijk en is gebaseerd op de overtuiging dat we toevallige grammaticale kenmerken alleen van algemene principes kunnen onderscheiden door zoveel mogelijk natuurlijke menselijke talen op zoveel mogelijk punten met elkaar te vergelijken. Deze overtuiging ligt aan de basis van de typologische aanpak in de taalkunde, waarvan Joseph Greenberg vaak als de moderne grondlegger beschouwd wordt. Een van de doelen van de taalkundige typologie is op basis van empirische gegevens over de talen van de wereld universele principes te formuleren. Er zijn in deze typologische traditie vier logisch mogelijke types universalia:

Type 1: Absolute, niet-implicationele universalia: *alle talen hebben eigenschap X*.

Type 2: Niet-implicationele tendensen: *de meeste talen hebben eigenschap X*.

Type 3: Absolute implicationele universalia: *alle talen die eigenschap X hebben, hebben ook eigenschap Y*

Type 4: Statistische implicationele universalia: *als een taal eigenschap X heeft, is het erg waarschijnlijk dat ze ook eigenschap Y heeft*

In een recent overzichtsartikel geven Nicholas Evans en Steven Levinson (2009) argumenten voor de bij typologen groeiende consensus dat universalia van het eerste type zo goed als onbestaande zijn. De kandidaat universalia van type 1 waarvan niet meteen tegenvoorbeelden gevonden kunnen worden, zijn ofwel triviaal, ofwel zo abstract dat ze nauwelijks falsifieerbaar zijn. Zelfs schijnbaar triviale universalia zoals “alle gesproken talen hebben lettergrepen” staan dankzij onze toegenomen kennis van de taalkundige variatie op de helling. De meest interessante en overtuigende “universalia” zijn de geconditioneerde tendensen van type 4. Hun geldigheid kan alleen (kwantitatief) getest worden in een uitgebreid en gebalanceerd staal van de talen van de wereld. Een gebalanceerd staal is nodig om genealogische verwantschap en onderlinge beïnvloeding door contact zoveel mogelijk te neutraliseren. Een uitgebreid staal is nodig om te achterhalen hoe sterk een tendens is.

In de rest van deze sectie zal ik bij wijze van illustratie kort bespreken wat een analyse van het Myene ons bijbrengt omtrent de complexe implicationele tendens (een universele van type 4, dus) die bekendstaat als de toegankelijkheidshiërarchie. Deze hiërarchie, oorspronkelijk geformuleerd door Edward Keenan en Bernard Comrie (1977), voorspelt onder andere welke elementen uit een zin toegankelijk zijn voor relativisering. De oorspronkelijke formulering van de hiërarchie bevatte teveel taalspecifieke elementen, zodat we hier met een enigszins aangepaste formulering moeten werken (1):

(1) S > P > G > T > Obl > Gen

In dit schema staat S voor *subject*, P (van *patient*) voor het object in een monotransitieve zin, G (van *goal*) grosso modo voor wat in de Nederlandse grammatica overeenstemt met het meewerkend voorwerp van een ditransitieve zin, T (van *theme*) voor het lijdend voorwerp in een ditransitieve zin, Obl voor *oblique* (vergelijkbaar met Nl. voorzetselvoorwerp), en Gen voor de bezitter in een genitiefconstructie. Met betrekking tot de vorming van relatieve bijzinnen voorspelt de hiërarchie onder andere het volgende (N.B. het gaat telkens om een sterke statistische tendens):

- Elke taal moet het subject van een zin kunnen relativiseren
- Elke strategie om relatiefzinnen te vormen moet van toepassing zijn op een aaneensluitend segment van de hiërarchie

Toegepast op het Nederlands zien we dat het subject van een zin inderdaad toegankelijk is voor relativisering. Op basis van een zin als *De regering is vorige week gevallen* kunnen we een nominale constituent maken met het subject van de zin als hoofd en de rest als dependens: *de regering die vorige week gevallen is*. De bezitter in een genitiefconstructie is eveneens toegankelijk voor relativisering. Vergelijk *De vaas van onze buurvrouw is op de grond gevallen* met *de buurvrouw van wie de vaas* (of archaïsch: *wier vaas*) *op de grond gevallen is*. Omdat in beide gevallen dezelfde strategie van relativisering gebruikt wordt (vooropplaatsing van het hoofd, gebruik van een relatiefpronomen), en beide elementen (S & Gen) zich op de twee uiteinden van de hiërarchie situeren, is de kans volgens de hierboven geformuleerde typologische veralgemening heel groot dat ook alle andere elementen op de hiërarchie toegankelijk zijn voor relativisering. Een strategie voor de vorming van relatiefzinnen moet immers beschikbaar zijn voor een aaneensluitend segment van de hiërarchie. Het Nederlands gedraagt zich wat dit betreft inderdaad net als de overgrote meerderheid (bijna honderd procent) van de talen van de wereld. Alle andere posities op de hiërarchie zijn toegankelijk voor relativisering (2).

- (2)
- a. P: de vaas die gisteren gevallen is
 - b. G: de student aan wie ik het toonsysteem van het Myene uitgelegd heb
 - c. T: het toonsysteem dat ik aan de student uitgelegd heb
 - d. Obl: de stokjes waarmee we onze lunch gegeten hebben

Het Myene, daarentegen vertoont een hoogst ongebruikelijk “gat” in de toegankelijkheidshiërarchie. Subjecten zijn toegankelijk voor relativisering door middel van een relatieve werkwoordsvorm. Het subject van de zin in (3) is het hoofd van de relatieve bijzin in (4). Merk op dat het formele verschil tussen de relatieve werkwoordsvorm in (4) en de niet-relatieve vorm in (3) puur tonaal is [1]. Alle Myenevoorbeelden in deze sectie zijn uitgelokt.

- (3) ó'ɲwánt àyòlín óy àbá

|ó'ɲw-ántò à-à-gòl-in-í ó-gà à-bá|

1-vrouw.DTP 1-RP-koop-APPL-RP 1-chef.DTP 6-mango.NTP

‘De vrouw kocht mango’s voor de chef.’

(4) ó’ɲwánt ‘áyòlìn óy àbá

|ó’ɲw-ántò á-à-yòl-in-ì ó-gà à-bá|

1-vrouw.DTP 1-RP-koop-APPL-RP.REL 1-chef.DTP 6-mango.NTP

‘de vrouw die mango’s kocht voor de chef’

Dezelfde strategie voor relativisering kan gebruikt worden voor voorzetselvoorwerpen (*Obl* in de hiërarchie), zoals *n òkwára* ‘met een hakmes’ in (5).

(5) òɲwá’n ónòmè àzèrín ónèndǵì ntínà y í’tótò n òkwára

‘De jongen heeft voor de leraar met een hakmes de bananenplant omgehakt.’

(6) òkwára wázèrìnì n òɲwá’n ónòmè ònèndǵì ntínà y í’tótò

‘het hakmes waarmee de jongen voor de leraar de bananenplant omgehakt heeft’

We verwachten dus dat ook alle tussenliggende posities op de toegankelijkheidshiërarchie met dezelfde strategie gerelativiseerd kunnen worden. Het object *nàyò* ‘huis’ van de zin in (7), bijvoorbeeld, d.i. P in de hiërarchie, zou het hoofd van een relatiefzin moeten kunnen zijn, die dezelfde structuur heeft als de relatiefzinnen in (4) en (6). Als we echter de relatiefzin in (8) analyseren, dan blijkt dat ze een passieve werkwoordsvorm heeft, gemarkeerd door het suffix *-o*. Met andere woorden, de P positie is niet rechtstreeks toegankelijk voor relativisering, maar moet eerst tot S-positie gepromoveerd worden door middel van passivering.

(7) myáyèní nàyò

|mí-à-yèn-í nágò|

1SG-RP-zie-RP 9.huis.DTP

‘Ik zag het huis.’

(8) náyò yáyèno myè

|nágò í-à-yèn-ò = myé|

9.huis.DTP IX-RP.zie-PASS.REL=1SG.PPR

‘het huis dat ik zag’ (letterlijk: het huis dat door mij gezien werd)

Dezelfde onrechtstreekse strategie is vereist voor de positie G en voor bepaalde types objecten in positie T. De “normale” strategie van relativisering, die gebruik maakt van een relatieve werkwoordsvorm en vooropplaatsing van het hoofd, bestrijkt dus een discontinu segment van de toegankelijkheidshiërarchie (onderstreept in het schema in (9)), tegen de sterke typologische tendens in.

(9) S > P > G > T_[+saillant] > T_[-saillant] > (Obl) > Gen

Het interessante is dat dit uiterst zeldzame minderheidspatroon de geldigheid van de toegankelijkheidshiërarchie kan bevestigen. (Merk op dat ik de term *minderheidspatroon* gebruik. Het heeft geen zin om van een uitzondering te spreken in het geval van een statistische universele.) Om het “gat” in de distributie van de relativiseringsstrategie te verklaren, kunnen we immers een beroep doen op deze hiërarchie. De redenering gaat als volgt. Het subject staat helemaal bovenaan de hiërarchie en is zo gemakkelijk toegankelijk voor relativisering dat alle talen het subject van een zin kunnen relativiseren (hierop zijn bij mijn weten nog geen uitzonderingen gevonden). Omdat het subject de positie is die bij uitstek gerelativiseerd wordt, kan er ooit een sterke tendens geweest zijn om andere posities in de zin eerst tot subject te promoveren door middel van passivering. De toegankelijkheidshiërarchie is echter niet alleen relevant voor de toegankelijkheid tot relativisering, maar ook voor de toegankelijkheid tot passivering. De tendens om te passiveren vooraleer te relativiseren zal dus sterker geweest zijn bovenaan de hiërarchie (P, G en, gedeeltelijk, T). Daar waar de tendens tot passivering het sterkst was, is ze geherinterpreteerd als verplichting, d.i. als grammaticale regel.

Het Myene bezorgt ons dus een kostbaar, want heel zeldzaam, minderheidspatroon met betrekking tot de toegankelijkheid tot relativisering van verschillende syntactische posities. Als ons scenario voor de totstandkoming van dit patroon klopt, bevestigt het Myene tezelfdertijd de geldigheid van de toegankelijkheidshiërarchie die ten grondslag ligt aan de typologische voorspellingen omtrent de toegankelijkheid tot relativisering. Dit toont aan dat de grammatica van elke taal verrassingen bevat die theoretisch relevant kunnen zijn en dat alle talen het verdienen grondig geanalyseerd en gedocumenteerd te worden, ook talen die behoren tot families waarvan we aannemen dat ze relatief goed bestudeerd zijn.

5. Conclusies

Of de documentaire taalkunde aanspraak kan maken op de status van aparte discipline binnen de taalkunde, zoals sommige beoefenaars beweren, heeft weinig belang. Belangrijk is de vraag of deze stroming erin zal slagen taalkundigen ervan te overtuigen kwaliteitsvolle primaire gegevens te verzamelen, ze op een duurzame manier te archiveren en ze ter beschikking te stellen aan de academische gemeenschap. De gewoonte van veldwerkers in de taalkunde (en wellicht ook in andere disciplines) om de gegevens waarop hun gepubliceerde analyses gebaseerd zijn niet openbaar te maken, is een discipline met wetenschappelijke ambities onwaardig. Deze lakse omgang met primaire gegevens is een rem op de cumulatieve vooruitgang van onze kennis, omdat elk onderzoek over een onderbestudeerde taal van nul moet herbeginnen met de verzameling van gegevens. Bovendien zorgt ze ervoor dat gepubliceerde analyses niet falsifieerbaar zijn en dat de relevantie van onderzoeksresultaten zelden verder strekt dan de interesses van de auteur op het moment van publicatie. Nochtans volstaat het een aantal goede gewoontes te creëren en vol te houden (zogenaamde “best practices”) om naast een beschrijving ook een solide documentatie te produceren in een onderzoeksproject dat op veldwerk steunt. We hebben een aantal van deze best practices overlopen in Sectie 3.

Het documentaire aspect van veldwerk is des te relevanter geworden nu de meerderheid van de talen van de wereld met uitsterven bedreigd is en nu taalkundigen er steeds meer van overtuigd raken dat de variatie tussen de talen van de wereld niet zo oppervlakkig is als de

meeste cognitieve wetenschappers in de tweede helft van de twintigste eeuw dachten (zie, opnieuw, Evans & Levinson 2009). We zullen het verschijnsel natuurlijke menselijke taal alleen begrijpen door zoveel mogelijk talen op zoveel mogelijk punten met elkaar te vergelijken. De studie van elke taal levert verrassingen op en vaak hebben die verrassingen de vorm van uitzonderingen op regels die voordien absoluut geacht werden. Het lijkt erop dat de positieve invloed van de documentaire taalkunde groeit en dat ze blijvend is in de gemeenschap van beschrijvende taalkundigen. De frustratie die overblijft, is dat we met veel te weinig veldwerkers zijn. Hoe zullen toekomstige generaties taalkundigen en cognitieve wetenschappers ons beoordelen als ze denken aan hoe we de resultaten van het millennialange natuurlijke experiment van taaldiversificatie hebben laten wegsmelten zonder ze te documenteren?

[1] Tonen worden als volgt genoteerd (geïllustreerd op de klinker a): à lage toon, á hoge toon, â dalende toon, ‘á verlaagde hoge toon. De volgende afkortingen worden gebruikt in de annotatie van voorbeelden: 1, 2, 3, ... nominaal klassenprefix; I, II, III, ... congruentieprefix; 1SG eerste persoon enkelvoud; APPL applicatiesuffix; DTP definitief toonpatroon; NTP neutraal toonpatroon; PASS passief suffix; PPR persoonlijk voornaamwoord; REL relatieve vorm van het werkwoord; RP ver verleden tijd.

Referenties

- Ambouroué, Odette. 2007. *Eléments de description de l’orungu, langue bantu du Gabon* (B11b). — Proefschrift. Brussel, Université Libre de Bruxelles.
- Evans, Nicholas & Levinson, Steven. 2009. The myth of language universals: Language diversity and its importance for cognitive science. — *Behavioral and Brain Sciences*, 32: 429-492.
- Grégoire, Claire. 1991. Premières observations sur le système tonal du myene-nkomi, langue bantoue du Gabon (B11e) — *Pholia*, 6: 107-129.
- Grégoire, Claire & Rekanga, Jean-Paul. 1994. Nouvelles hypothèses diachroniques sur la cl 10b du myene-nkomi (B11e). — *Africana Linguistica*, 11: 71-77.
- Himmelman, Nikolaus (1998). Documentary and descriptive linguistics. — *Linguistics*, 36:161-195.
- Keenan, Edward & Comrie, Bernard. 1977. Noun phrase accessibility and universal grammar. — *Linguistic Inquiry* 8: 63-99.
- Raponda Walker, André. 1934. *Dictionnaire mpongwé-français*. — Metz: Libre Lorraine.
- Mouguiama-Daouda, Patrick. 1990. Esquisse d'une phonologie diachronique du mpongwè. — *Pholia*, 5: 114-120.
- Philippon, Gérard & Puech, Gilbert (ongepubliceerd). Tonal domains in Galwa.