

Hoe goed begrijpen Nederlandstaligen en Afrikaanstaligen geschreven Fries?¹

Charlotte Gooskens (Rijksuniversiteit Groningen) en Renée van Bezooijen (Radboud Universiteit Nijmegen)

1 Inleiding

Het klinkt aannemelijk dat talen die nauw verwant zijn, en dus veel overeenkomsten vertonen op het gebied van de fonologie, het lexicon en de grammatica, tot op zekere hoogte onderling te verstaan zijn, ook zonder enige voorkennis. De geldigheid van deze aanname is bijvoorbeeld aangetoond voor de drie Noord-Germaanse talen Noors, Zweeds en Deens (Haugen 1966, Maurud 1976, Börestam Uhlmann 1994, Zeevaert 2004), voor de Romaanse talen Spaans en Portugees (Jensen 1989) en voor de Slavische talen Slowaaks en Tsjechisch (Budovičová 1987). In deze bijdrage zullen we ons richten op de onderlinge verstaanbaarheid van de drie West-Germaanse talen Fries, Nederlands en Afrikaans, met de nadruk op de schriftelijke communicatie. Concreet betekent dit dat we de resultaten zullen presenteren van een experimentele studie waarin we Nederlandstaligen en Afrikaanstaligen die geen kennis hadden van het Fries, Friese teksten hebben voorgelegd waaruit we bepaalde woorden hadden verwijderd. Als maat voor het tekstbegrip hebben we geteld hoeveel woorden de proefpersonen op de correcte plaats in de tekst wisten terug te plaatsen. Daarnaast hebben we de Friese teksten linguïstisch geanalyseerd en gerelateerd aan de talen van de lezers. Op deze manier hebben we een antwoord gezocht op de volgende twee vragen:

1. Hoe goed begrijpen Nederlandstaligen en Afrikaanstaligen geschreven Fries?
2. Kunnen eventuele verschillen in tekstbegrip worden verklaard uit verschillen in linguïstische relaties tussen het Nederlands en het Fries en tussen het Afrikaans en het Fries?

Voor zover wij weten, is dit de eerste keer dat de begrijpelijkheid van het Fries is onderzocht voor sprekers van het Afrikaans. De begrijpelijkheid van het Fries voor sprekers van het Nederlands is eerder bestudeerd in een serie experimenten door Van Bezooijen & Van den Berg (1999a, 1999b, 1999c, 2000). Hierbij ging het echter om gesproken communicatie en niet om schriftelijke communicatie. Daarnaast is er één onderzoek uitgevoerd naar de begrijpelijkheid van het geschreven Fries voor Nederlandstaligen (Van Bezooijen & Gooskens 2005). In paragraaf 2 zullen we dit eerdere onderzoek naar de verstaanbaarheid van het Fries voor Nederlandstaligen samenvatten. In paragraaf 3 zullen we kort ingaan op de historische verwantschap tussen het Fries, Nederlands en Afrikaans. In paragraaf 4 beschrijven we de methode en resultaten voor tekstbegrip en in paragraaf 5 de methode en resultaten voor de linguïstische analyse. In paragraaf 6, ten slotte, wordt het artikel afgesloten met een conclusie.

2 Eerder onderzoek naar de verstaanbaarheid van het Fries voor Nederlandstaligen

In het eerste experimentele onderzoek naar de verstaanbaarheid van het Fries voor Nederlandstaligen (Van Bezooijen & Van den Berg 1999a, 1999b) is een vergelijking getrokken tussen de verstaanbaarheid van het gesproken Fries en drie dialecten van het Nederlands, namelijk het Bedums (Groningen), het Valkenburgs (Limburg) en het Tielts (West-Vlaanderen) in een informatie-rijke context. Het spraakmateriaal bestond uit beschrijvingen van vijf plaatjes waarop alledaagse situaties (interieur van een keuken,

winkelstraat, etc.) waren afgebeeld. De beschrijvingen, die afkomstig waren van telkens drie ter plekke geboren en opgegroeide vrouwen, werden ten behoeve van het onderzoek opgebroken in korte fragmenten. Deze werden in hun oorspronkelijke volgorde auditief aan luisteraars (mensen zonder kennis van het Fries of dialecten van het Nederlands) aangeboden. Ze kregen tevens een uitgeschreven versie van de fragmenten, waarin alle woorden in het Nederlands waren voorvertaald behalve de zelfstandige naamwoorden. Deze dienden de luisteraars zelf op de daarvoor in de tekst opengelaten plaatsen in het Nederlands in te vullen. De percentages correct vertaalde zelfstandige naamwoorden waren 86% voor het Fries, 94% voor het Gronings, 86% voor het Limburgs en 64% voor het West-Vlaams. In dit onderzoek nam het Fries dus een middenpositie in, samen met het Limburgs.

In een vervolgonderzoek (Van Bezooijen & Van den Berg 1999c) werd de relatieve verstaanbaarheid van dezelfde vier variëteiten onderzocht, maar nu in een informatie-arme context. Dat wil zeggen dat de hoeveelheid informatie die de luisteraars bij het vertalen van de zelfstandige naamwoorden ter beschikking stond op verschillende manieren was ingeperkt. Ten eerste bevatten de fragmenten in het eerdere onderzoek vaak meerdere zelfstandige naamwoorden, bijvoeglijke naamwoorden, bijwoorden of werkwoorden. Als de luisteraar een van die woorden correct verstond, hielp dat bij het verstaan van de andere woorden. In het tweede onderzoek zijn alleen fragmenten aangeboden met een enkel zelfstandig naamwoord. Ten tweede werden in het eerdere onderzoek de fragmenten aangeboden in dezelfde volgorde waarin ze door een bepaalde spreekster bij het beschrijven van een bepaald plaatje waren uitgesproken. Bij het vertalen van een zelfstandig naamwoord konden de luisteraars dus gebruik maken van alles wat ze in de voorafgaande fragmenten hadden gehoord. In het tweede onderzoek werden de fragmenten weliswaar nog steeds bloksgewijs per variëteit aangeboden (dus bijvoorbeeld eerst alle Friese fragmenten, dan alle Limburgse fragmenten, etc.), maar binnen elk blok in een willekeurige volgorde, met de plaatjes en spreeksters door elkaar. In het tweede onderzoek waren de luisteraars dus vrijwel geheel aangewezen op een correcte interpretatie van de akoestische informatie vervat in het bevroegde zelfstandige naamwoord zelf, zonder hulp uit de context. Het Fries bleek nu veel meer problemen op te leveren voor de verstaanbaarheid (58% correct) dan het Gronings (94% correct) en het Limburgs (80% correct) en evenveel als het West-Vlaams (58% correct).

Het spraakmateriaal dat in het eerste en tweede onderzoek werd gebruikt, bestond uit eenvoudige, korte, beschrijvende zinnestelsels met betrekking tot een enkel semantisch domein (huis-, tuin- en keukentermen). Het deed hiermee geen recht aan de grammaticale en lexicale rijkdom van het Fries. Dit was de reden om nog een derde verstaanbaarheidsonderzoek op te zetten, met complexer spraakmateriaal (Van Bezooijen & Van den Berg 2000). Hierbij werd gebruik gemaakt van een luistervaardigheidstoets van de Afûk, het opleidingsinstituut voor het Fries. Deze toets maakte in 1995 deel uit van het examen Fries voor niet-Friestaligen. Verstaanbaarheid werd vastgesteld door middel van 22 meerkeuzevragen bij een vrij lange, voorgelezen tekst. In ons onderzoek heeft een groep Friestaligen als referentie gediend voor de prestaties van een groep Nederlandstaligen. De verstaanbaarheid van samenhangende, voorgelezen teksten bleek significant lager voor Nederlandstaligen dan voor Friestaligen, maar het verschil was klein (gemiddelde aantallen correcte antwoorden van respectievelijk 17,4 en 20,5). Wij wijten de goede prestaties van de Nederlandstalige luisteraars aan aspecten van de gebruikte methode: de zorgvuldigheid van de uitspraak van de aangeboden tekst, de informatie over de inhoud van de tekst die kon worden afgeleid uit de (in het Nederlands gestelde) meerkeuzevragen en het feit dat alle Nederlandstalige luisteraars een taalkundige achtergrond hadden.

De drie beschreven onderzoeken hadden alle betrekking op de verstaanbaarheid van gesproken Fries. Het enige onderzoek dat is uitgevoerd naar de begrijpelijkheid van het geschreven Fries voor Nederlandstaligen zonder kennis van het Fries, is uitgevoerd door Van

Bezooijen & Gooskens (2005). In dit onderzoek is een vergelijking getrokken tussen de begrijpelijkheid van geschreven Fries en geschreven Afrikaans. Er werd gebruik gemaakt van een variant van de cloze-test, waarbij woorden uit de tekst waren gehaald en door de proefpersonen correct dienden te worden teruggeplaatst. Nederlandstalige taalstudenten bleken relatief weinig moeite te hebben om deze taak uit te voeren voor het Afrikaans (gemiddeld 82% correct). Het Fries leverde aanzienlijk meer problemen op (gemiddeld 50% correct). Attitudes leken geen rol te spelen. Er waren wel een aantal linguïstische maten die een verklaring zouden kunnen vormen voor het verschil in moeilijkheidsgraad. Zo bevatten de Friese teksten meer woorden die niet verwant waren met de corresponderende woorden in het Nederlands dan de Afrikaanse teksten.

Het hier gepresenteerde onderzoek is qua opzet te vergelijken met Van Bezooijen & Gooskens (2005), er is gebruik gemaakt van dezelfde cloze-test. Er zijn echter drie verschillpunten.

1. In het huidige onderzoek is geen gebruik gemaakt van taalstudenten maar van middelbare scholieren in het voorlaatste jaar van hun opleiding. We hoopten zo de generaliseerbaarheid van de resultaten te vergroten.
2. In tegenstelling tot het eerdere onderzoek bevat het huidige onderzoek een referentieconditie, zodat de prestaties voor de onbekende taal konden worden vergeleken met de prestaties voor de eigen taal.
3. In het huidige onderzoek is geen vergelijking getrokken tussen de prestaties van Nederlandstaligen voor het Fries vergeleken met het Afrikaans, maar tussen de prestaties van Nederlandstaligen vergeleken met Afrikaanstaligen voor het Fries.

Het huidige onderzoek bevatte net zoals Van Bezooijen & Gooskens (2005) zowel een attitudecomponent als een linguïstische component ter verklaring van het tekstbegrip. We beperken ons hier echter tot de resultaten voor de linguïstische maten. Voor we hier op ingaan, geven we eerst wat achtergrondinformatie over de verhoudingen tussen de drie betrokken talen.

3 Historische verwantschap tussen Fries, Nederlands en Afrikaans

Het Fries en Afrikaans zijn beide nauw verwant aan het Nederlands, maar de historische achtergronden van deze relaties zijn verschillend. Grofweg kan men stellen dat het Fries en het Nederlands op een bepaald moment twee talen waren die steeds meer naar elkaar toe zijn gegroeid, terwijl het Afrikaans uit het Nederlands is ontstaan en zich daar steeds verder van heeft verwijderd.

Het Fries behoort tot de West-Germaanse tak van de Germaanse talen. Historisch gezien is het Fries nauwer verwant met het Engels dan met het Nederlands. In de loop van de tijd zijn het Fries en het Engels echter meer en meer uit elkaar gegroeid en het Fries en het Nederlands steeds meer naar elkaar toe gegroeid. Het Fries werd eeuwenlang vrijwel alleen gesproken en niet geschreven, het had een lage status als 'boerentaal'. In 1844 werd het *Frysk Selskip foar Tael- en Skriftekennisse* opgericht, het begin van een Friese beweging die zich vooral ten doel stelde om het Fries tot volwaardige cultuurtaal uit te bouwen. Hoewel het Fries nu een geschreven en, iets minder duidelijk, een gesproken standaard heeft, wordt het nog steeds sterk door het Nederlands beïnvloed (Ytsma 1995).

Het Afrikaans is oorspronkelijk een dialect dat zich ontwikkelde onder uit Nederland afkomstige kolonisten die zich op de Kaap hadden gevestigd. Daarom lijkt het Afrikaans veel op het Nederlands van de 17e eeuw. In de loop van de tijd veranderde het Afrikaans steeds meer onder invloed van onder andere het Standaardnederlands en het taalgebruik van

anderstaligen die het Nederlands onvoldoende beheersten. In het begin schreef men vooral in het Nederlands. In de 19e eeuw begon men voor het eerst de Afrikaanse taal op te schrijven. In 1875 werd het *Genootschap van de Regte Afrikaners* opgericht. Zij streden voor de erkenning van het Afrikaans als cultuurtaal. Tot in de 20^e eeuw werd het Afrikaans als een Nederlands dialect beschouwd, maar sinds 1921 heeft het de officiële status van een zelfstandige taal.

4 Tekstbegrip

Het eerste deel van het onderzoek betrof een test waarmee de begrijpelijkheid van het geschreven Fries werd vastgesteld voor Nederlandstalige en Afrikaanstalige lezers die het Fries niet kenden.

4.1 Methode

Proefpersonen

De Nederlandstalige proefpersonen waren vwo-eindexamenleerlingen in hun voorlaatste jaar van middelbare scholen uit Zuid-Drente en Overijssel in de buurt van Zwolle en Coevorden. Vooraf moesten ze op een vragenformulier aangeven hoe vaak ze Fries spraken en schreven (actieve kennis) en lezen en hoorden (passieve kennis). We hebben alleen de gegevens gebruikt van proefpersonen die verklaarden nooit Fries te spreken of te schrijven. Verder werden ook proefpersonen uitgesloten van deelname die verklaarden vaker dan één keer per jaar een van de twee talen te lezen of te horen. Zo werden 34 leerlingen (16 jongens en 18 meisjes) geselecteerd die geen actieve en zo goed als geen passieve kennis hadden van het Fries. Zij hadden een gemiddelde leeftijd van 16,6 jaar. Ze hadden allen het Nederlands als moedertaal. Twee proefpersonen spraken thuis naast Nederlands ook Duits.

De Afrikaanstalige proefpersonen waren leerlingen van scholen waar het Afrikaans als voertaal wordt gebruikt. Ze kwamen uit Henneman in de provincie Vrystaat en uit Douglas in de provincie Noord-Kaap. Het Zuid-Afrikaanse schoolsysteem verschilt van het Nederlandse systeem. Alle Zuid-Afrikaanse kinderen gaan naar dezelfde soort middelbare school, maar de verschillende vakken worden op twee niveaus aangeboden. We gebruikten slechts gegevens van proefpersonen die minstens vier vakken op het hoogste niveau volgden, hetgeen toegang geeft tot de universiteit. Op deze manier konden we ervan uitgaan dat het opleidingsniveau van de Zuid-Afrikaanse proefpersonen vergelijkbaar was met dat van de Nederlandse proefpersonen. Geen van de Zuid-Afrikaanse proefpersonen had actieve of passieve kennis van het Fries. In totaal voldeden 37 Zuid-Afrikaanse leerlingen (16 jongens en 21 meisjes) aan de criteria. Hun gemiddelde leeftijd was 17,2 jaar. Hun moedertaal was Afrikaans en daarnaast spraken drie leerlingen thuis ook Engels.

Begripstest

Als basis voor de begripstest gebruikten we twee Nederlandse krantenartikelen van gemiddelde moeilijkheidsgraad.² De ene tekst ('feministen-tekst') gaat over het beeld van de vrouw dat jongeren via de moderne muziekzenders krijgen voorgespiegeld en bevat 317 woorden. De andere tekst ('relatie-tekst') gaat over datingbureaus op Internet en bevat 329 woorden. Het tekstbegrip van de proefpersonen is getoetst door middel van een variant van de cloze-test.³ Hiertoe zijn in beide teksten vijf bijvoeglijke naamwoorden, vijf bijwoorden, vijf zelfstandige naamwoorden en vijf zelfstandige werkwoorden geselecteerd. In principe was dit ieder tiende woord in de tekst, behalve als dit woord niet in een van de vier woordcategorieën viel. Dan namen we het volgende woord in de tekst. De teksten moesten twee keer volgens

deze procedure worden doorgewerkt om voor iedere categorie voldoende items te krijgen. De twintig woorden werden in alfabetische volgorde boven de tekst gezet en in de tekst vervangen door strepen. De twee Nederlandse teksten werden vervolgens in het Afrikaans en het Fries vertaald.⁴ Daarna werden dezelfde woorden als in de Nederlandse teksten uit de teksten verwijderd. In Appendix A wordt als voorbeeld de Friese relatie-test gegeven. Het was de taak van de proefpersonen de twintig woorden in de originele context terug te plaatsen. De proefpersonen kregen hiervoor precies 10 minuten de tijd. Uit vooronderzoek was gebleken dat vrijwel iedereen de taak binnen deze tijd kan volbrengen. Het aantal correct teruggeplaatste woorden is als maat genomen voor het tekstbegrip.

Design

Eerst kregen de proefpersonen een lijst met persoonlijke vragen, met betrekking tot herkomst, thuistaal, leeftijd en geslacht. Vervolgens moesten ze een aantal vragen beantwoorden over het contact met de betreffende taal. Deze informatie is gebruikt om de proefpersonen te selecteren (zie boven). Daarna volgden de begripstesten. Alle proefpersonen kregen eerst een Friese tekst (L2) en daarna een tekst in hun eigen taal (L1, Nederlands of Afrikaans). De ene helft van de proefpersonen kreeg de feministen-tekst in L2 en de relatie-tekst in L1. De andere helft van de proefpersonen kregen de relatie-tekst in L2 en de feministen-tekst in L1. Zo werd een systematisch leereffect tegengegaan.

4.2 Resultaten

Om de eerste onderzoeksvraag “Hoe goed begrijpen Nederlandstaligen en Afrikaanstaligen geschreven Fries?” te kunnen beantwoorden, hebben we het aantal woorden geteld dat correct in de originele zinscontext is teruggeplaatst. Ten behoeve van een grotere stabiliteit hebben we de gegevens voor de twee teksten bij elkaar opgeteld. De resultaten worden gepresenteerd in Tabel 1.

We richten onze aandacht eerst op de prestaties voor de eigen taal. We zien in Tabel 1 dat de Nederlandse proefpersonen hier geen problemen mee hadden; zij wisten de cloze-test voor het Nederlands vrijwel foutloos uit te voeren (97,1% correcte responsies). Voor de Zuid-Afrikaanse proefpersonen was de taak echter moeilijker dan we hadden verwacht (68,0% correct). Het matige resultaat voor de Zuid-Afrikaanse leerlingen kan niet worden verklaard uit het feit dat hun Afrikaanse taalvaardigheid laag is. Het Afrikaans is hun moedertaal, ze leven in een overwegend Afrikaanstalige omgeving en ze gaan naar Afrikaanstalige scholen. Drie Zuid-Afrikaanse proefpersonen waren wel degelijk in staat het maximum aantal punten (20) te halen in de test en vijf anderen haalden bijna het maximum met 18 punten. Het is onduidelijk waarom niet ook de andere leerlingen de taak naar behoren hebben kunnen volbrengen. De Nederlandse en Zuid-Afrikaanse proefpersonen waren zorgvuldig gematcht wat betreft hun opleiding en leeftijd (zie paragraaf 4.1). Misschien waren de onderwerpen van de teksten minder interessant of vertrouwd voor de Zuid-Afrikaanse leerlingen dan voor de Nederlandse leerlingen. Of misschien hadden de Zuid-Afrikaanse leerlingen minder ervaring met (dit soort) testen dan de Nederlandse leerlingen. Deze verklaringen zijn echter niet aannemelijk, aangezien er geen verband was tussen de prestaties voor L1 en L2. Als testervaring of belangstelling een rol hadden gespeeld, had men een vergelijkbaar effect verwacht voor de twee talen.

Als we de moeilijk te interpreteren prestaties van de Zuid-Afrikaanse leerlingen voor hun eigen taal even terzijde schuiven, zijn er in onze ogen twee indicaties dat het Fries moeilijker was voor de Zuid-Afrikaanse leerlingen dan voor de Nederlandse leerlingen. Ten eerste kan dit worden afgeleid uit het absolute percentage correcte responsies. Het gemiddelde percentage correcte responsies voor de Nederlandse proefpersonen was 31,9. Dit betekent dat

ze erin slaagden om gemiddeld meer dan 6 van de 20 items correct te lokaliseren. Voor de Zuid-Afrikaanse proefpersonen bedroeg het gemiddelde percentage correct 6,7%. Dit betekent dat ze gemiddeld nog geen twee items correct in de tekst wisten te plaatsen. Gemiddeld over de twee teksten is er een verschil van 25,2% tussen de prestaties van de Nederlandse en de Zuid-Afrikaanse leerlingen. Een tweede indicatie dat de Zuid-Afrikaanse leerlingen meer moeite hadden met het Fries dan de Nederlandse leerlingen zijn de prestaties van de leerlingen van wie het zeker is dat ze geen enkele moeite hadden met de taak op zich. De acht Zuid-Afrikaanse leerlingen die hoog scoorden op de L1-test (minstens 18 van de 20 items correct) hadden allemaal lage scores (niet mee dan vier items correct) op de L2-test. Dus hoewel deze leerlingen in principe prima in staat waren om de taak uit te voeren, lukte hun dit niet voor het Fries. Geen enkele Zuid-Afrikaanse leerling had meer dan vier items correct. Bij de Nederlandse leerlingen lag dit geheel anders. Er waren vier Nederlandse leerlingen die zowel voor hun eigen taal als het Fries vrijwel alle (minstens 18 van de 20) items correct in de tekst konden plaatsen.

Tekst	Proefpersonen				t-test		
	Nederlands		Zuid-Afrikaans		t	df	p
	Gemid.	Range	Gemid.	Range			
Fries	31,9	1-16	6,7	0-4	8,172	69	.000
Eigen taal	97,1	12-20	68,0	6-20	7,114	69	.000

Tabel 1. Links: gemiddeld percentage correcte responsies en range van de absolute aantallen correct van de twee groepen proefpersonen. Rechts: bijbehorende resultaten van een t-test voor onafhankelijke steekproeven.

5 Linguïstische maten

5.1 Methode

Om zicht te krijgen op de problemen waarmee de proefpersonen in de cloze-test geconfronteerd werden, hebben we zeven linguïstische maten berekend. Als eerste stap in de linguïstische analyse hebben we de complete Nederlandse en Afrikaanse teksten opgelijnd met de Friese teksten. Oplijnen houdt in dat we zo goed mogelijk de corresponderende woorden in de twee talen hebben gepaard, met soms een kleine aanpassing in de woordvolgorde. Zo zijn de elementen van de dubbele ontkenning in het Afrikaans - *nie* na de persoonsvorm en *nie* aan het eind van de zin - samengevoegd om aan te sluiten bij het Friese *net* 'niet'. Nadat de woorden waren opgelijnd, hebben we voor ieder Fries-Nederlands en Fries-Afrikaans woordpaar de aard van de relatie tussen de corresponderende leden vastgesteld. De maten A, B en C hebben te maken met hun onderlinge verwantschap. Maat D bestaat uit de Levenshtein-afstand, die de mate van gelijkenis tussen de woordvormen uitdrukt. De maten E, F en G, ten slotte, drukken de transparantie van de relatie van de Friese woorden met de corresponderende woorden in de eigen taal uit. Bij het berekenen van de verschillende maten hebben we persoonsnamen, die per definitie identiek zijn in de verschillende talen, buiten beschouwing gelaten. In de verdere tekst verwijst de term 'bronwoord' naar het woord in de oorspronkelijke tekst in de talen van de lezer (Nederlands en Afrikaans). De term 'doelwoord' verwijst naar het corresponderende woord in het Fries dat de lezer moet proberen te begrijpen.

Maat A. Percentage verwante woorden, direct of indirect via een synoniem

In zijn algemeenheid kan worden aangenomen dat talen die veel verwante woorden ('cognates') bevatten, onderling makkelijker te begrijpen zijn dan talen die veel niet-verwante woorden ('non-cognates') bevatten. In onze teksten is directe verwantschap tussen doelwoord en bronwoord in principe wel bevorderend, maar niet essentieel voor de afleidbaarheid van de betekenis van het doelwoord. De betekenis van het Friese of Afrikaanse doelwoord kan soms even goed worden afgeleid aan de hand van de gelijkenis met een synoniem van het bronwoord. Bijvoorbeeld: in de Friese tekst staat *maatskippij* terwijl het bronwoord in de oorspronkelijke Nederlandse tekst *samenleving* is. Hoewel *maatskippij* en *samenleving* historisch niet verwant zijn, kan de betekenis van het Friese *maatskippij* toch correct door de Nederlandse lezer worden afgeleid via het Nederlandse woord *maatschappij*, dat een synoniem is van *samenleving*. De Nederlandse lezer heeft geen weet van de oorspronkelijke Nederlandse tekst, dus voor hem is het verschil tussen bronwoord en synoniem van bronwoord niet relevant.

De eerste linguïstische maat die we hebben berekend, is het percentage direct of via een synoniem verwante woorden. Omdat we aannamen dat verwante inhoudswoorden (zelfstandige naamwoorden, zelfstandige werkwoorden, bijvoeglijke naamwoorden, telwoorden) meer bijdragen tot tekstbegrip dan verwante functiewoorden (lidwoorden, voegwoorden, voorzetsels, voornaamwoorden, koppelwerkwoorden, hulpwerkwoorden, partikels, bijwoorden), hebben we de percentages ook apart voor deze twee woordcategorieën berekend.

Maat B. Percentage via een paradigma verwante woorden

Er is nog een tweede manier waarop de betekenis van het doelwoord kan worden afgeleid zonder dat er sprake hoeft te zijn van een directe verwantschap met het bronwoord, namelijk via een paradigma. Dit is een verzamelterm voor alle gevallen waarin het Friese doelwoord gelijkenis vertoont met een woord in het Nederlands of het Afrikaans dat via de grammatica gerelateerd is aan het Nederlandse of Afrikaanse bronwoord. Bijvoorbeeld: in de Friese tekst staat *binne* corresponderend met het Nederlandse *zijn* (derde persoon meervoud). De vormen *binne* en *zijn* zijn niet direct verwant aan elkaar en zijn evenmin via een synoniem aan elkaar gerelateerd, tenminste niet in de standaardtaal. Toch is er een relatie via de Nederlandse woorden *ben* of *bent*, twee enkelvoudsvormen uit het paradigma van het werkwoord *zijn*. Het is mogelijk dat de Nederlandse lezers deze relatie tussen het Friese *binne* en het Nederlandse *ben(t)* doorzien. De lexicale betekenis van het woord komen zij daarmee wel op het spoor, alleen raken ze wellicht in de war over de grammaticale betekenis van het woord, denken ze dat het om een enkelvoudsvorm gaat in plaats van om een meervoudsvorm. De paradigmatische relatie hoeft niet via het Nederlandse bronwoord zelf te gaan, het kan even goed via een synoniem van dit woord. Zoals gezegd, voor de lezer maakt het niet uit, die ziet de oorspronkelijke en vertaalde tekst niet naast elkaar.

Het percentage via een paradigma verwante woorden (direct of via een synoniem) vormt de tweede linguïstische maat. Het gaat in vrijwel alle gevallen om functiewoorden.

Maat C. Percentage niet-verwante woorden

Naast de direct of indirect (via synoniem of paradigma) verwante woorden zijn er ook gevallen waarin er geen enkele relatie is tussen doelwoord en bronwoord. Het betreft dan gevallen waarin het Friese woord geen enkele gelijkenis vertoont met een woord met een vergelijkbare betekenis in het Nederlands of het Afrikaans. Voorbeelden: Fr. *gleonens* vs. Ned. *wellust* en Fr. *jonge 'jongen'* vs. Afr. *ou*. We laten hierbij een eventuele verwantschap van het Friese woord met een derde taal, zoals het Engels of Duits, buiten beschouwing. Ten eerste weten we niet hoe goed de talenkennis van de proefpersonen is en ten tweede weten we niet in hoeverre de strategie om ook nog andere talen in het interpretatieproces te gebruiken

wordt toegepast. Het percentage niet-verwante woorden is het complement van maat 1 en 2 (de drie maten tellen op tot 100%), maar voor de volledigheid hebben we het hier toch als een aparte maat opgenomen, ook weer opgesplitst naar functiewoorden en inhoudswoorden.

Maat D. Levenshtein-afstand

Het is onmogelijk de betekenis af te leiden van een Fries doelwoord dat geen enkele vormelijke relatie vertoont met een Nederlands of Afrikaans woord met een gerelateerde betekenis (maat C). In principe is dit wel mogelijk bij verwante woorden (maat A en B). De mate waarin de lezer kan profiteren van het feit dat doelwoord en bronwoord direct dan wel via een synoniem of een paradigma aan elkaar verwant zijn, kan echter variëren. Woorden kunnen in de loop van de tijd zo veranderd zijn, dat de woordvormen nog maar weinig gemeenschappelijke kenmerken hebben. Met de Levenshteinmaat kan de mate van gelijkheid tussen woordvormen automatisch en dus objectief worden bepaald. De Levenshtein-afstand is hier eenvoudig gedefinieerd als het minimum aantal *letters* dat vervangen, ingevoegd of verwijderd moet worden om het ene woord om te zetten in het andere woord.⁵ Er wordt rekening gehouden met de woordlengte door de totale kosten te delen door het aantal ophijningen binnen het woord. De methode wordt uitgebreid besproken in Heeringa (2004). Hier volgt een voorbeeld van het berekenen van de Levenshtein-afstand tussen het Nederlandse woord *zoeken* en het corresponderende Friese woord *sykje*:

	Ophijningen						
	1	2	3	4	5	6	7
Nederlands	z	o	e	k	-	e	n
Fries	s	y	-	k	j	e	-
kosten	1	1	1		1		1

De totale kosten ($1+1+1+1+1=5$) worden gedeeld door het aantal ophijningen (7). Dit levert een afstand op van 0,71 of 71%. De totale afstand tussen de twee teksten is dan de gemiddelde afstand over de afstanden tussen alle woordparen. De Levenshtein-afstanden zijn bepaald op basis van de teksten waarin in voorkomende gevallen het oorspronkelijke bronwoord is vervangen door een synoniem (zie de uitleg bij maat A). Met andere woorden: de Levenshtein-afstand is niet bepaald tussen het Friese *maatskippij* en het oorspronkelijke Nederlandse *samenleving*, maar tussen het Friese *maatskippij* en het Nederlandse *maatschappij*, een synoniem voor *samenleving*. Als er voor een niet-verwant woord geen synoniem voorhanden was, is de Levenshtein-afstand bepaald tussen het originele bronwoord en het doelwoord.

Maat E. Doorzichtigheid van de lexicale betekenis

Met de Levenshtein-afstanden wordt slechts vastgesteld in hoeverre letters gelijk zijn of verschillend. De maat houdt geen rekening met de communicatieve consequenties van de verschillen. Het is bijvoorbeeld gebleken dat verschillen aan het begin van een woord een grotere belemmering vormen bij woordherkenning dan verschillen aan het einde (Broerse en Zwaan 1966). Op dit moment is het nog niet mogelijk met het effect van deze en andere factoren, zoals de relatieve bijdrage van klinkers en medeklinkers, rekening te houden bij het berekenen van de Levenshtein-afstanden. Om deze reden hebben we zelf de doorzichtigheid van de betekenissen ingeschat, waarbij we onze oordelen baseerden op onze eigen intuïties als taalgebruikers. We hebben vier gradaties van doorzichtigheid van de lexicale relatie onderscheiden. Deze gradaties worden hieronder nader uitgewerkt.

- Lexicale relatie geheel doorzichtig (0 punt)

De twee verwante woorden hebben synchroon dezelfde vorm, wat betekent dat het doorzien van de lexicale relatie tussen bronwoord en doelwoord voor de lezer geen enkel probleem vormt. Het Friese is identiek, qua vorm én betekenis, aan het Nederlandse of Afrikaanse woord. Het woord *sterk* is een voorbeeld voor het Fries en het Nederlands en het woord *geweld* voor het Fries en Afrikaans. Soms hebben we ook woorden met een klein spellingsverschil een score van 0 punten gegeven. Voorwaarde was dat het spellingsverschil niet mag leiden tot een uitspraakverschil bij het hardop lezen van het woord. Voorbeelden: Fr. *buro's* vs. Ned. *bureaus*, Fr. *eksplosyf* vs. Afr. *eksplosief*.

- Lexicale relatie redelijk doorzichtig (1 punt)

De twee verwante woordvormen vertonen synchroon een grote gelijkenis, waardoor de relatie van het Friese woord met het Nederlandse of Afrikaanse woord redelijk makkelijk te doorzien is. Meestal verschillen doelwoord en bronwoord slechts in één enkele letter. Voorbeelden: Fries *strang* vs. Ned. *streng*, Fr. *frou* vs. Afr. *vrou* 'vrouw'.

- Lexicale relatie behoorlijk ondoorzichtig (2 punt)

De twee verwante woordvormen vertonen synchroon maar een beperkte gelijkenis, waardoor de relatie tussen doelwoord en bronwoord moeilijk te doorzien is. Vrijwel altijd is er sprake van meerdere letters verschil. Voorbeelden: Fr. *jierren* vs. Ned. *jaren*, Fr. *ieuwenâlde* vs. Afr. *eeu-oue* 'eeuwenoude'.

- Lexicale relatie geheel ondoorzichtig (3 punt)

De twee verwante woordvormen vertonen synchroon zo weinig gelijkenis dat het voor de Nederlandse of Zuid-Afrikaanse lezer vrijwel onmogelijk is de betekenis van het Friese woord af te leiden uit de overeenkomst met het corresponderende Nederlandse of Afrikaanse woord. Meestal is er sprake van een verschil in het merendeel van de letters. Voorbeelden: Fr. *deis* vs. Ned. *dagelijks*, Fr. *jierren* vs. Afr. *jare* 'jaren'.

De scoring van de doorzichtigheid van de lexicale relatie van de Friese woorden met de Nederlandse woorden is uitgevoerd door de tweede auteur. Aangezien het subjectieve element hier groter is dan bij de andere maten, is de scoring in dit geval ook nog een keer, onafhankelijk van de eerste, uitgevoerd door een tweede persoon. Het percentage identieke scores was 82,2%. De correlatie tussen de scores was hoog (Pearson's $r = .92$). Op grond van deze uitkomsten hebben we besloten dat de scores zoals gegeven door de eerste auteur voldoende betrouwbaar waren om ongewijzigd te worden gehandhaafd. Om de doorzichtigheidscores voor Zuid-Afrikaanse lezers te bepalen, werkte de tweede auteur samen met een moedertaalspreker van het Afrikaans.

Maat E wordt berekend door per taal alle punten op te tellen en te middelen over het totale aantal woordparen. De scores kunnen variëren tussen 0 en 3 en zijn genormeerd tot percentages door ze te delen door 3 en ze dan te vermenigvuldigen met 100.

Maat F. Doorzichtigheid van de grammaticale relatie

Naast een correcte interpretatie van de lexicale betekenis is ook een correcte interpretatie van de grammaticale betekenis van een woord noodzakelijk voor een goed tekstbegrip: (a) Is het duidelijk om wat voor woordsoort het gaat? (b) Is het duidelijk om welke tijd, getal, geslacht, persoon, etc. het gaat? Een voorbeeld van het eerste type probleem is het Friese *net* 'niet' dat door zowel Zuid-Afrikaanse als Nederlandse lezers als 'net' kan worden geïnterpreteerd met een geheel andere betekenis ('zojuist' in het Nederlands, 'alleen, precies' in het Afrikaans). Een voorbeeld van het tweede probleem is het Friese *sy* 'zij', dat Afrikaanstaligen als een enkelvoudvorm zullen interpreteren, terwijl het in het Fries (en Nederlands) ook compatibel is

met het meervoud. Informatie over de grammaticale betekenis is belangrijk om de structuur van de tekst te doorzien: de relatie tussen de woorden binnen de zin en de relatie tussen de zinnen binnen de tekst. Het ontbreken van grammaticale informatie kan net zo goed als het ontbreken van lexicale informatie nadelige gevolgen hebben voor de interpretatie van tekst en dus een belasting vormen voor het tekstbegrip. De doorzichtigheid van de grammaticale relatie tussen doelwoord en bronwoord werd apart van de lexicale relatie gecodeerd. Over het algemeen is er sprake van een tweedeling: de grammaticale relatie is wel doorzichtig (0 punt) of hij is niet doorzichtig (1 punt). We hebben de twee typen grammaticale afleidbaarheid (a) en (b) apart gecodeerd, omdat ze in onze ogen beide op een eigen manier verwarring kunnen veroorzaken. De scores kunnen daarom variëren tussen 0 en 2 en zijn genormeerd tot percentages door ze te delen door 2 en dan te vermenigvuldigen met 100.

Maat G. Doorzichtigheid van de complete relatie

Om de totale doorzichtigheid van de relatie tussen doelwoord en bronwoord te bepalen hebben we de punten voor de problemen met de afleidbaarheid van de lexicale betekenis (maat E) en die voor de problemen met de grammaticale betekenis (maat F) bij elkaar opgeteld. Omdat maat E kan lopen van 0 tot 3 en maat F van 0 tot 2, wordt maat E zwaarder gewogen dan maat F. Dit komt in onze ogen overeen met de werkelijke bijdrage van de lexicale en de grammaticale betekenis aan tekstbegrip. De waarden kunnen liggen tussen 0 tot 5 en zijn genormeerd tot percentages door ze te delen door 5 en dan te vermenigvuldigen met 100.

5.2 Resultaten

In Tabel 2 staan de uitkomsten met betrekking tot de eerste drie maten voor de linguïstische relatie tussen het Fries en het Nederlands en tussen het Fries en het Afrikaans: percentage direct of indirect via een synoniem verwante woorden (maat A), percentage via een paradigma verwante woorden (maat B) en percentage niet-verwante woorden (maat C). De percentages worden ook apart gepresenteerd voor de functie- en inhoudswoorden.

We zien dat het percentage direct of via een synoniem verwante woorden veel groter is voor de Nederlandse proefpersonen (93,8%) dan voor de Zuid-Afrikaanse proefpersonen (79,2%). Dit verschil geldt zowel voor functiewoorden als voor inhoudswoorden. Ook in categorie B, woorden die verwant zijn via een paradigma, is het percentage hoger voor de Nederlandse lezers (4,6% voor functiewoorden) dan voor de Zuid-Afrikaanse lezers (2,3%). Er is ook een groot verschil in het percentage niet-verwante woorden (3,7% voor Nederlanders en 19,5% voor Zuid-Afrikanen). Het effect van het verschil in categorie B op de begrijpelijkheid is niet bij voorbaat duidelijk. Waar in ieder geval wel een sterk effect op het tekstbegrip te verwachten is, zijn de percentages in categorie A en C. Veel verwante en weinig niet-verwante woorden, zoals in het geval van de Nederlandse lezers, maakt een tekst makkelijk interpreteerbaar. Voor de Zuid-Afrikaanse lezers is tekstinterpretatie veel moeilijker; zij worden geconfronteerd met weinig verwante en veel niet-verwante woorden.

	Nederlands			Afrikaans		
	functie	inhoud	totaal	functie	inhoud	totaal
A direct verwant of via synoniem	93,4	94,1	93,8	80,1	78,1	79,2
B verwant via paradigma	4,6	-	2,5	2,3	-	1,3
C niet-verwant	2,0	5,9	3,7	17,6	21,9	19,5

Tabel 2. Maat A, B en C. Percentage met het Fries verwante woorden in de Nederlandse en Afrikaanse teksten, voor functiewoorden, inhoudswoorden en totaal

Het is echter niet voldoende om te kijken naar het percentage woorden die vallen in de verschillende categorieën van verwantschap. De woorden in categorie C zijn per definitie oninterpreteerbaar en daarom kunnen we ervan uitgaan dat ze problemen opleveren voor de tekstbegrip. Bij de woorden in categorie A en B ligt dit anders. Deze woorden zijn weliswaar historisch verwant met woorden in de taal van de proefpersonen, maar de verwantschap hoeft niet altijd af te leiden zijn uit de overeenkomsten tussen de hedendaagse woordvormen. Daarom hebben we ook linguïstische afstanden gemeten tussen het Afrikaans en het Nederlands ten opzichte van het Fries met behulp van de Levenshtein-maat (maat D). De uitkomsten zijn te zien in Tabel 3. We zien ten eerste dat de niet-verwante woorden (maat C) zeer hoge waarden, dus veel verschillen, hebben voor beide talen. Dit kan ook niet anders, als er al letters overeenkomen dan berust dit op toeval. Hetzelfde geldt in iets mindere mate voor woorden die met elkaar verwant zijn via een paradigma (maat B). Vaak hebben deze een andere vorm (vergelijk Fr. *wy* 'wij' en Afr. *ons*), maar het komt ook geregeld voor dat er enige overeenkomst is tussen de woorden (vergelijk de *h* in Fr. *harren* en Ned. *hun*). Het interessantst zijn de direct of via een synoniem verwante woorden (maat A). De Levenshtein-afstand is hier gemiddeld een stuk kleiner, wat men ook zou verwachten. Immers, deze categorie omvat naast afwijkende vormen ook veel identieke vormen. De afstanden tussen de functiewoorden is hetzelfde voor het Afrikaans en het Nederlands, maar voor de inhoudswoorden is er een groot verschil (30,1% voor Nederlanders en 40,2 % voor Zuid-Afrikanen). Een *t*-test voor onafhankelijke steekproeven liet zien dat dit verschil significant is op het 0,01 niveau ($t = 5,2481$, $df = 496$). Dit geeft dus aan dat in de categorie direct of via een synoniem verwante inhoudswoorden de vorm van het geschreven Afrikaanse woord meer afwijkt van het Fries dan die van het geschreven Nederlandse woord.

	Nederlands			Afrikaans		
	functie	inhoud	totaal	functie	inhoud	totaal
A direct verwant of via synoniem	37,6	30,1	34,2	37,6	40,2	38,8
B verwant via paradigma	67,1	-	67,1	40,6	-	40,6
C niet-verwant	83,9	78,2	79,9	79,6	78,5	79,0

Tabel 3. Maat D. Gemiddelde Levenshtein-afstand (in %) tussen Fries-Nederlandse en Fries-Afrikaanse woordparen voor functiewoorden, inhoudswoorden en totaal, apart voor de maten A, B en C.

In Tabel 4 laten we de resultaten van de doorzichtigheidsmaten zien voor de Nederlandse en Afrikaanse woorden die verwant zijn met de Friese woorden (maat A en B). Er wordt een onderscheid gemaakt tussen doorzichtigheid van de lexicale relatie (maat E), doorzichtigheid van de grammaticale relatie (maat F) en doorzichtigheid van de complete relatie (maat G). We presenteren de scores in hun oorspronkelijke vorm maar ook omgerekend in percentages om het makkelijker te maken de drie maten te vergelijken. De afleidbaarheid van de lexicale betekenis van functiewoorden voor Nederlandse lezers is bijvoorbeeld $(1,24/3)*100 = 41,3\%$.

Uit Tabel 4 kan worden opgemaakt dat de lexicale relatie van de Friese woorden met de Afrikaanse woorden (1,63) moeilijker te doorzien is dan de lexicale relatie met de Nederlandse woorden (1,33). Het verschil is het grootst bij inhoudswoorden (1,92 voor Nederlanders vs. 1,45 voor Zuid-Afrikanen). Dezelfde tendens vinden we voor de doorzichtigheid van de grammaticale relatie, hoewel de verschillen hier kleiner zijn (0,05 voor Nederlanders vs. 0,09 voor Zuid-Afrikanen). Al met al lijkt vooral maat D een goed inzicht te geven in de oorzaken van de lagere begripsscores bij de Zuid-Afrikanen dan bij de Nederlanders: bij verwante woorden is het doorzien van de lexicale relatie moeilijker voor Zuid-Afrikanen dan voor Nederlanders.

Doorzichtigheid	Nederlands			Afrikaans		
	functie	inhoud	totaal	functie	inhoud	totaal
E Lexicale relatie (range 0-3)	1,24 (41,3)	1,45 (48,3)	1,33 (44,3)	1,39 (46,3)	1,92 (64,0)	1,63 (54,3)
F Grammaticale relatie (range 0-2)	0,03 (1,5)	0,07 (3,5)	0,05 (2,5)	0,13 (6,5)	0,04 (2,0)	0,09 (4,5)
G Complete relatie (range 0-5)	1,28 (22,7)	1,52 (30,4)	1,39 (27,8)	1,52 (30,4)	1,96 (39,2)	1,71 (34,2)

Tabel 4. Maat E, F en G. Gemiddelde doorzichtigheid van de verwantschapsrelatie van Friese woorden met het Nederlands en Afrikaans (maat A en B), uitgesplitst naar functiewoorden, inhoudswoorden en totaal. Hoe hoger het getal, hoe ondoorzichtiger de relatie en hoe groter de belasting voor het tekstbegrip. Tussen haakjes zijn de scores omgerekend in percentages.

We zien dus dat zowel bij de Levenshtein-afstand (maat D) als bij de doorzichtigheidsscores (maat E, F, en G) de afstand van het Fries tot het Afrikaans groter is dan tot het Nederlands. Over alle woorden samen is de correlatie tussen de Levenshtein-afstanden en de scores voor de doorzichtigheid van de complete relatie (maat F) .84 voor het Afrikaans en .67 voor het Nederlands. De percentages verklaarde variantie zijn respectievelijk 70.6% en 44.9%. Naast (flink wat) overeenkomsten zijn er dus ook verschillen tussen de twee maten. Dit komt doordat bij de Levenshtein-afstanden alle verschillen even zwaar tellen, terwijl bij de doorzichtigheidsscores rekening wordt gehouden met het communicatieve effect van de aard van de correspondenties. Zo krijgen sommige woordparen een doorzichtigheidsscore van 0, terwijl er wel degelijk sprake is van een spellingsverschil (zie voorbeelden in 3.1), wat bij de Levenshteinmaat tot een afstand leidt.

6 Conclusie

In dit onderzoek hebben we vastgesteld hoeveel Nederlandstalige en Afrikaanstalige lezers begrijpen van geschreven Friese teksten van een gemiddelde moeilijkheidsgraad. De twee

groepen proefpersonen waren op drie relevante punten met elkaar gematcht: leeftijd (ongeveer 17 jaar), opleidingsniveau (voorbereidend wetenschappelijk) en kennis van het Fries (geen actieve kennis en verwaarloosbaar weinig passieve kennis). Het tekstbegrip is gemeten door middel van een variant van de cloze-test. De Nederlandstaligen bleken minder moeite te hebben de Friese teksten te begrijpen dan de Afrikaanstaligen.

Teneinde het verschil tussen de prestaties van de Nederlandstalige en Afrikaanstalige lezers te verklaren, hebben we zeven linguïstische maten berekend. Deze waren alle gebaseerd op een paarsgewijze Fries-Afrikaanse en Fries-Nederlandse vergelijking van corresponderende woorden in de teksten. Vrijwel alle maten leverden resultaten op die wezen op een grotere linguïstische afstand tussen het Fries en het Afrikaans dan tussen het Fries en het Nederlands. Ten eerste bevatten de Nederlandse teksten meer aan het Fries verwante woorden en minder niet aan het Fries verwante woorden dan de Afrikaanse teksten. Ten tweede waren de objectieve vormverschillen tussen Fries-Nederlands verwante woorden (geoperationaliseerd als het aantal verschillende letters) kleiner dan tussen Fries-Afrikaans verwante woorden. Ten derde zijn de relaties tussen de Fries-Nederlandse verwante woorden makkelijker te doorzien dan de relaties tussen de Fries-Afrikaanse verwante woorden. De uitkomsten van de linguïstische analyse zijn dus over de hele lijn in overeenstemming met de uitkomsten van de begripstest: Friese teksten zijn makkelijker te begrijpen voor Nederlandstalige lezers dan voor Afrikaanstalige lezers omdat de linguïstische afstand tussen het Fries en het Nederlands kleiner is dan tussen het Fries en het Afrikaans.

Historisch gezien ligt het ook voor de hand dat het Afrikaans meer zou verschillen van het Fries dan het Nederlands. Het Afrikaans is sinds het zich aan het begin van de 17e eeuw van het Nederlands heeft afgesplitst steeds verder van het Nederlands gedivergeerd, op alle linguïstische niveaus. Het fonologische en het morfologische systeem hebben op veel punten reductie ondergaan. Ook is er sprake van divergentie in het lexicon. Het Afrikaans heeft recente leenwoorden in het Nederlands niet opgenomen en aan de andere kant ontbreken in het Nederlands weer woorden die het Afrikaans (in beperkte mate) aan andere talen heeft ontleend of zelf heeft gecreëerd (neologismen). Naarmate het Afrikaans verder van het Nederlands is gaan afwijken, is automatisch ook de afstand tot het Fries vergroot. Het Afrikaans wijkt daardoor nu meer af van het Fries dan het Nederlands. Het gaat hierbij om de relatieve afstand tot het Fries. In absolute zin is de afstand tussen het Fries en het Nederlands de afgelopen decennia verkleind als gevolg van interferentie van het Nederlands in het Fries. Het Fries convergeert met het Nederlands vooral op het gebied van het lexicon.

Literatuur

Bezooijen, R. van & Berg, R. van den (1999a)

‘Verstaanbaarheid van het Gronings, Fries, Limburgs en West-Vlaams: Waar zitten de problemen’, *Artikelen van de Derde Sociolinguïstische Conferentie*, Lunteren, 49-60.

Bezooijen, R. van & Berg, R. van den (1999b)

‘Taalvariëteiten in Nederland en Vlaanderen: hoe staat het met hun verstaanbaarheid?’, *Taal en Tongval 51*, 15-33.

Bezooijen, R. van & Berg, R. van den (1999c)

‘Word intelligibility of language varieties in the Netherlands and Flanders under minimal conditions’, in: R. van Bezooijen & R. Kager (reds.) *Linguistics in the Netherlands*, 16, 1-12.

Bezooijen, R. van & Berg, R. van den (2000)

‘Hoe verstaanbaar is het Fries voor niet-Friestaligen?’, *Filologia Frisica*. Ljouwert: Fryske Akademy, 9-26.

- Bezooijen, R. van & Gooskens, C. (2005)
 ‘Intertalig tekstbegrip. De begrijpelijkheid van Friese en Afrikaanse teksten voor Nederlandse lezers.’ *Nederlandse Taalkunde*, 10, 129-152.
- Björnsson, C.H. (1968)
Läsbarhet [Leesbaarheid]. Stockholm: Liber.
- Börestam Uhlmann, U. (1994)
Skandinaver samtaler. Språkliga och interaktionella strategier i samtal mellem danskar, norrmän och svenskar [Gesprekken tussen Scandinaviërs. Talige en interactionele strategieën in gesprekken tussen Denen, Noren en Zweden]. Uppsala: Uppsala Universiteit.
- Broerse, A.C. & Zwaan E.J. (1966)
 ‘The information value of initial letters in the identification of words.’ *Journal of Verbal Learning and Verbal Behavior*, 5, 441-446.
- Budovičová, V. (1987)
 ‘Literary languages in contact (A sociolinguistic approach to the relation between Slovak & Czech today).’ In: J. Chloupek, J. Nekvapil et al. (reds.), *Reader in Czech sociolinguistics*. Amsterdam/Philadelphia: John Benjamins, 156-175.
- Haugen, E. (1966)
 ‘Semicommunication: The language gap in Scandinavia.’ *Sociological inquiry*, 36, 280-297.
- Heeringa, W. (2004)
Measuring dialect pronunciation differences using Levenshtein distances. Groningen: Groningen dissertations in linguistics (Grodil).
- Jensen, J. (1989)
 ‘On the mutual intelligibility of Spanish and Portuguese.’ *Hispania*, 72, 849-852.
- Maurud, Ø. (1976)
Nabospråksförståelse i Skandinavia. En undersøkelse om gjensidig forståelse av tale- og skriftspråk i Danmark, Norge og Sverige [Onderlinge verstaanbaarheid van talen in Scandinavië. Een onderzoek naar wederzijdse begrip van geschreven en gesproken taal in Denemarken, Noorwegen en Zweden]. Stockholm: Nordiska rådet.
- Ytsma, J. (1995)
Frisian as first and second language. Ljouwert: Fryske Akademy.
- Zeevaert, L. (2004)
Interskandinavische Kommunikation. Strategien zur Etablierung von Verständigung zwischen Skandinaviern im Diskurs. Hamburg: Dr. Kovač.

Appendix A

De Friese relatie-test

Hier staan 20 woorden. Vul deze in in de onderstaande Friese tekst.

al	dwaan	eksplosyf	fantastysk	ferhalen
hawwe	knappe	lang	leit	lestige
misbrûk	moatte	nea	pear	soms
teloarsteld	ûnbetroubere	wurden	wurket	ynternet

Net lang mei ljocht hier mar in lytse, keale sukkel

Jierliks sykje tsientûzenen frijgesellen fia in relaasjebemiddelingsburo of op ----- nei in partner. ----- mei súkses, faak ek net. It docht bliken dat ----- manlju yn it echt âld en ûnsjoch binne, en ----- buru's binne nearne te finen sadree't it ynskriuwjild oermakke is.

It klikte ----- ûnder it chatten. De jonge mei wa't Annemarie op ynternet yn kontakt kommen wie, foldie oan al har easken: hy wie ----- en sportyf en hie ljocht hier. Skreau er. "Mar by it earste ôfspraakje yn it kafee siet der in sukkel op my te wachtsjen", sei de 27-jierrige frou. "Hy wie lyts, keal en seach der ek nochris tsien jier âlder út as de leeftyd dy't er op 'e side opjûn hie." Annemarie wiksele trije ----- mei de man en gong liik en teloarsteld fuort. Dat mearke wie út.

Foar sokke klanten is Joke Pronk oansteld. Sy ----- foar de Algemiene Ferein Relaasjeburo's (AFR). In ----- saak. "Dy minsken binne ----- ", neffens Pronk, "Se hawwe faak neat oan it werom beteljen fan in ----- hûndert euro, se wolle benammen stoom ôfblaze." Pronk ----- fêst hokker buru's harren ôfspraken net neikommen binne en dielt 'giele kaarten' út.

It tal bemiddelingsburo's is de ôfrûne jierren - benammen op ynternet - ----- groeid. Der binne der ----- fjouwer- oant fiifhûndert. Mei de groei nimt ek it ----- ta. It docht bliken dat in soad ynternetburo's ferdwûn binne sadree't it ynskriuwjild oermakke is. "Ynienen besteane dy siden net mear. Of se ----- in kantoar yn Kazachstan." Pronk is op 'e hichte fan de -----, mar dy dogge harren neffens har net foar by de fyftjin buru's mei in AFR-karmerk, dêr't sy foar wurket. Sy riedt minsken oan -----, harren eigen e-mailadres te jaan, mar allinnich te mailen fia de side. "As minsken ôfprate wolle, advisearje wy mei klam dat te ----- op in iepenbier plak. It binne wier net allinnich psychopaten dy't op ynternet in frou sykje, sa't guon sizze, mar jo ----- wol foarsichtich wêze.

Noten

¹ Wij danken Vincent van Heuven, Sulette Bruwer en Bart Alewijnse voor hun hulp in verschillende stadia van het onderzoek.

² Wij hebben de leesbaarheid van de teksten bepaald met behulp van de zogenaamde LIX-index (Björnsson 1968). Deze bestaat uit het gemiddelde aantal woorden per zin plus het percentage woorden die langer zijn dan zeven letters. Teksten met een LIX-waarde van 35 tot 44 hebben een gemiddelde moeilijkheidsgraad. De gemiddelde LIX-waarde voor de twee Nederlandse teksten was 42.

³ De cloze-test is in 1953 in de Verenigde Staten ontwikkeld door William Taylor en sinds die tijd veel gebruikt voor het meten van tekstbegrip. Soms worden de woorden boven de tekst geplaatst, zoals in dit onderzoek, maar soms moeten de proefpersonen zelf bedenken welke woorden het best in de tekst passen.

⁴ De Friese tekst is in eerste instantie vertaald door een Friestalige student Fries en vervolgens gecontroleerd door een medewerker van de Fryske Akademy. De Afrikaanse tekst is vertaald door een Afrikaanstalige student Engels en gecontroleerd door een Afrikaanstalige taalkundige.

⁵ We hebben dus gekozen voor afstanden op basis van binaire verschillen: verschillende symbolen dragen bij tot een linguïstische afstand, identieke symbolen niet. Voor gesproken teksten wordt vaak een meer geavanceerde versie gebruikt, waarbij rekening wordt gehouden met het feit dat de afstand tussen sommige klanken groter is dan tussen andere.