

Taalvariatie en sociale media

SND



taal:
unie



Stichting
Cultureel
Erfgoed
Zeeland

Taalvariatie en sociale media

Lezingen gehouden op de
internationale streektaalconferentie in
Middelburg op 7 oktober 2016

Redactie:

Veronique De Tier, Ton van de Wijngaard
& Anne-Sophie Ghyselen

De internationale streektaalconferentie van 7 oktober 2016 in Middelburg werd georganiseerd door de Stichting Nederlandse Dialecten (SND), Variaties vzw. Koepelorganisatie voor dialecten en oraal erfgoed in Vlaanderen en de Stichting Cultureel Erfgoed Zeeland (SCEZ). We danken de Nederlandse Taalunie voor de financiële steun.

© Stichting Nederlandse Dialecten, Leiden 2017

ISBN 978-90-73869-00-4

Niets uit deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm, internet of op welke wijze ook zonder voorafgaande schriftelijke toestemming van de uitgever.

De uitgever heeft ernaar gestreefd de rechten van de illustraties volgens de wettelijke bepalingen te regelen. Wie desondanks meent rechten te kunnen doen gelden, kan zich alsnog tot de uitgever richten.

Druk: Nevelland Graphics

Opmaak: Anne-Sophie Ghyselen

Kaft: vectortekening door Allan McAvoy

Een uitgave van de Stichting Nederlandse Dialecten

Inhoud

Voorwoord	7
<i>Veronique De Tier</i>	
Regionale identiteitsconstructie en taalvariatie op sociale media Limburg als casus	13
<i>Leonie Cornips</i>	
Brabants dialect op Facebook en Twitter	25
<i>Jos Swanenberg</i>	
Het gebruik van Friese varianten op sociale media	33
<i>Lysbeth Jongbloed-Faber</i>	
#dialect	52
Culturele modellen en de perceptie van taalvariëteiten op Twitter	
<i>Karliën Franco en Laura Rosseel</i>	
Isoglossen in Twitter? Op zoek naar tweetgeografie	53
<i>Hans van Halteren en Roeland van Hout</i>	
Hun twitteren	65
Tweets als bron voor onderzoek naar syntactische taalvariatie	
<i>Stefan Grondelaers, Roeland van Hout en Hans van Halteren</i>	

Voorwoord

In 2016 organiseerde de Stichting Nederlandse Dialecten (SND) haar jaarlijkse streektaalconferentie in het Zeeuwse Middelburg. De Stichting Cultureel Erfgoed Zeeland was zo vriendelijk in te staan voor het onthaal, en dat was niet de eerste keer. Zij deden dat eerder al, samen met Erfgoed Brabant bij de derde streektaalconferentie (over 'dialect van schoot tot school'), die in 2008 plaatsvond in Rilland. Die conferentie was een vervolg op eerdere bijeenkomsten waarin streektaalbeleid centraal stond, zoals de edities 'Lang leve de dialecten. Streektaalbeleid in Nederland' (Maastricht, 2006) en 'Streektaalbeleid en duurzaamheid' (Noordwolde, 2007). Precies tien jaar zit er tussen de geboorte van de eerste streektaalconferentie en die van 7 oktober 2016, die gehouden werd in de Week van het Nederlands. De streektaalconferenties hebben ondertussen een stevige traditie.

Vanaf de vierde editie in 2009 raakte de SND betrokken bij het initiatief, samen met Variaties vzw, de Vlaamse koepelorganisatie voor dialecten en oraal erfgoed. Plaats van gebeuren was toen Brussel. In de loop van de volgende jaren organiseerde de SND de conferenties met diverse thema's in vrijwel alle regio's in Nederland en België: 'Streektaal in de zorg' (Enschede, 2010), 'Streektaal als erfgoed?!' (Elst, 2012), 'Digi-streek-taal' (Hasselt, 2013), 'Streektaalliteratuur' (Hoogeveen, 2014) en 'Taalvariatie in een grootstedelijke omgeving' (Brussel, 2015). Niet van elke streektaalconferentie zijn de lezingen in druk verschenen, maar op de website van de Stichting (www.nederlandsedialecten.org) kunt u de oudere publicaties digitaal inkijken, samen met enkele niet gepubliceerde lezingen.

Het thema van de streektaalconferentie in 2016 was taalvariatie en sociale media. Sociale media zijn relatief nieuw, en lijken op het eerste gezicht wat in tegenstrijd met dialect, dat vaak als oubollig wordt gezien. Toch blijkt er op sociale media heel wat te bewegen in dialect en streektaal. Een aantal gevestigde waarden die zich al jaren bezighouden met de taal in sociale media, zoals Reinhild Vandeckerckhove en Benny De Decker van de Universiteit Antwerpen en Dong Nguyen en Dolf Trieschnigg van de Universiteit Twente, konden niet aanwezig

zijn op de conferentie. Er bleven echter zeker voldoende bevlogen onderzoekers over om een hele dag te vullen met lezingen over dit moderne thema.

Wereldwijd heeft Twitter meer dan 200 miljoen gebruikers; het aantal Nederlandse twitteraars wordt geschat tussen de een en twee miljoen. Alleen de Nederlandse twitteraars zorgen samen al voor een dagelijkse stroom van zo'n vier miljoen tweets. Antal van den Bosch verzamelde samen met Erik Tjong Kim Sang en Evert Lammerts een grote hoeveelheid van deze tweets. De teller staat al boven de vijf terabytes en groeit nog altijd aan. De vele twitterberichten van het onderzoeksproject TwiNL vormen dus een verzameling big data waaruit interessante trends en conclusies kunnen worden gehaald. Het selecteren van de tweets is een tweestapsproces. In eerste instantie wordt gekeken of er een bekend Nederlands woord in voorkomt, of een hashtag die in ons land vaak wordt gebruikt. In tweede instantie wordt met een automatische taalherkenner een verdere selectie gemaakt. Daarnaast worden de tweets van de 5000 meest productieve twitteraars in Nederland standaard toegevoegd aan de dataverzameling. Om Twitterberichten van het internet te plukken, gebruiken de onderzoekers een in Zweden ontwikkeld hulpmiddel. Om analyses uit te voeren, maken ze gebruik van een Hadoop-installatie die staat opgesteld bij SURFsara in Amsterdam. SURFsara zorgt ook voor de opslagcapaciteit die nodig is om alle tweets te kunnen bewaren, met als resultaat dat de onderzoekers in meer Twitterberichten kunnen zoeken dan bij Twitter zelf. Dat zoeken kan op de website <http://twiqs.nl>, waarbij een zoekopdracht afhankelijk van de belasting van de website tussen de 10 seconden en 1 minuut in beslag neemt. De resultaten van de zoekopdracht kunnen in grafiekvorm of in de vorm van een landkaart weergegeven worden. Antal Van den Bosch, hoogleraar taal- en spraaktechnologie aan de Radboud Universiteit, directeur van het Centre for Language Studies aldaar en kort na de conferentie ook directeur van het Meertens Instituut, demonstreerde tijdens de streektaalconferentie Twiqs.nl en liet zien hoe dialectwoorden machinaal geogost kunnen worden op Twitter. U kunt er ook thuis mee aan de slag.

In Vlaanderen bestuderen vooral Reinhold Vandekerckhove, oud-voorzitter van de Stichting Nederlandse Dialecten en haar medewerker Benny De Decker al een aantal jaren de chattaal in Vlaanderen. Maar ook aan de VUB is er aandacht voor. Ester Magis bereidt er bijvoorbeeld een proefschrift voor over chat- en tussentaal in Vlaanderen en bracht hierover een verslag op de conferentie, waarin zij de specifieke kenmerken van chattaal in Vlaanderen toelichtte. Vlaamse tiesners blijken diverse variëteiten van het Nederlands te gebruiken, zoals Standaardnederlands, maar ook tussentaal, de informele en spontane spreektaal van

Vlamingen. Net zoals de gesproken tussentaal in Vlaanderen verschilt van regio tot regio en zelfs van spreker tot spreker, is ook het online taalgebruik sterk onderhevig aan regionale en idiolectische variatie. Slechts een handvol zogenaamd typische kenmerken blijkt door tieners uit heel Vlaanderen frequent gebruikt te worden: *gij*, *ge* en *u(w)* in plaats van *jij* en *je*, het weglaten van de eind-*t* in woordjes als *dat*, *wat* en *niet*, het diminutiefsuffix *-ke* (*rokske* in plaats van *rokje*) en de toevoeging van *dat* na onderschikkende voegwoorden, zoals in de zin 'ik weet wie dat er komt'. Dialect blijkt minder gebruikt te worden door de chat-tende Vlaamse tieners.

Zowel in Nederland als in Vlaanderen wordt dus volop onderzoek gedaan naar de taal op sociale media, zoals in chatberichten en in tweets. Enkele Nederlandse en Vlaamse onderzoekers lichtten hun onderzoek toe op deze elfde streektaalconferentie.

Leonie Cornips begon in het uiterste zuidoosten van Nederland in Limburg. Zij is onderzoeker sociolinguïstiek en syntaxis aan het Meertens Instituut (KNAW) in Amsterdam en hoogleraar Taalcultuur in Limburg aan de Universiteit van Maastricht. Zij deed onder meer onderzoek naar het gesproken Nederlands in Heerlen, stedelijke jongerentaal (ook wel straattaal genoemd) en nieuwe vormen van Nederlands zoals Surinaams-Nederlands. Zij leidt ook onderzoek naar dialect en sociale media in Limburg en Friesland. Leonie belicht het Limburgs.

Jos Swanenberg is behalve bestuurslid van de SND ook adviseur streektalen bij Erfgoed Brabant en bijzonder hoogleraar Diversiteit in taal en cultuur bij het departement Cultuurstudies van Tilburg University. Hij doet onder meer onderzoek naar taal, cultuur en identiteit van jonge en nieuwe Brabanders, en laten net die bezig zijn met sociale media. Hij bekeek het thema taal en sociale media met een Noord-Brabantse bril.

Na haar studies in Leeuwarden en aan de Rijksuniversiteit Groningen heeft Lysbeth Jongbloet eerst tien jaar in het bedrijfsleven gewerkt, maar sinds 2013 is ze werkzaam bij de Fryske Akademy, waar ze het gebruik van Fries op sociale media onderzoekt. In november 2015 is ze begonnen aan een promotietraject op de Fryske Akademy bij de Universiteit Maastricht, met Leonie Cornips als promotor. Hoofdvraag van haar onderzoek is 'Hoe kan het gebruik van Friese varianten op sociale media worden verklaard?'

Roeland van Hout is geen onbekende op de streektaalconferenties. Hij was de

initiatiefnemer van de streektaalconferenties, hoogleraar Toegepaste taalwetenschap en Variatielinguïstiek aan de Radboud Universiteit. Samen met zijn collega Hans van Halteren zoekt hij op een iets andere manier – via Twitter – naar de oude Limburgse isoglossen.

Beide heren werkten ook samen met Stef Grondelaers (Radboud Universiteit) aan een lezing over *hun hebben*. Deze subjectsvorm wekt in Nederland nog altijd grote irritatie op, maar de wetenschappelijke aandacht voor het verschijnsel is niet proportioneel met de heisa die het genereert. Dat komt omdat Nederlanders steeds vaker *hun hebben* zeggen, maar het vooralsnog niet schrijven, en taalkundigen hun analyses doorgaans op geschreven taalmateriaal baseren. Grondelaers, Van Hout & Van Halteren demonstreren in hun lezing dat tweets een prima data-alternatief vormen voor onderzoek naar syntactische variatie: ze bevatten immers veel *hun*-constructies en laten in weerwil van hun korte formaat toe om alle beïnvloedende factoren bij de constructie in beeld te brengen. Veel meer nog dan ‘gewoon’ tekstmateriaal vormen tweets bovendien een uitstekende biotoop om de sociale betekenis van *hun* te bestuderen.

Karliën Franco en Laura Rosseel van de KU Leuven lichten het gebruik van de hashtag ‘#dialect’ op twitter toe. Ze bestuderen het gebruik en de populariteit van dialecthashtag op Twitter en trachten zo te achterhalen wat tweets ons leren over de culturele modellen die een rol spelen bij het concept dialect en of sociale media gebruikt kan worden om culturele modellen van taal(-variatie) te bestuderen (de centrale onderzoeksvraag). In het algemeen blijkt dat de meerderheid van de #dialect-twitteraars aansluiten bij een romantische ideologie over taalvariatie, waarbij taalvariatie de expressie van verschillende identiteiten toelaat. Slechts een minderheid beschouwt dialectvariatie als hinderlijk (rationeel). Ze concluderen dat Twitter een interessante databron kan vormen, maar dat er wel belangrijke beperkingen zijn, zoals de beknoptheid van tweets en de waarschijnlijke bias in de groep gebruikers.

De streektaalconferentie bood een mooi overzicht van wat er op het vlak van taal en sociale media zoal te vinden en te beleven is in Nederland en Vlaanderen; we hopen u in deze bundel te kunnen laten meeproeven van het bestaande onderzoek. Het valt te verwachten dat de wetenschappelijke interesse voor het verband tussen taal en sociale media de volgende periode exponentieel zal uitbreiden; digitale media zullen immers naar alle waarschijnlijkheid een steeds belangrijker maatschappelijke positie innemen. Wij kijken alvast uit naar de resultaten van dat onderzoek.

De streektaalconferentie van 2016 werd georganiseerd in samenwerking met Variaties vzw, Koepelorganisatie voor dialecten en oraal erfgoed in Vlaanderen, met steun van de Nederlandse Taalunie en in het kader van de Week van het Nederlands. Ik dank de Nederlandse Taalunie voor het vertrouwen dat zij in de Stichting Nederlandse Dialecten hebben en voor de financiële steun hiervoor. Zonder die steun zou het onmogelijk zijn om elk jaar een boeiende conferentie te organiseren. Verder dank ik ook mijn medebestuurders van de SND en de instituten waar zij werken voor hun enthousiasme om ondanks de volle agenda's toch iedere keer naar de voorbereidende vergaderingen te komen. Zonder sprekers is een conferentie onmogelijk. Zij verdienen ook een welgemeend dankjewel voor de inzet en de moeite om hun lezing om te zetten in een leesbare tekst. En ten slotte dank ik ook graag mijn mederedacteuren die hun schaarse tijd vrijwillig hebben opgeofferd om deze bundel met lezingen samen te stellen en na te lezen. Extra dank gaat nog eens naar Anne-Sophie die er met haar lay-outvaardigheden in geslaagd is een leesbare bundel aan te bieden.

Veronique De Tier
Voorzitter Stichting Nederlandse dialecten
September 2017

Regionale identiteitsconstructie en taalvariatie op sociale media.

Limburg als casus

Leonie Cornips

Twitteren in een minderheidstaal

*@X kiek ns ff ofste powerpoint oaf hobs, maach iech murge daan ff get beij diech oaf-
maken :p aubbb xx aandars kreijg onvoldoende :p x*

Bovenstaande tweet (vertaald 'kijk eens even of je de powerpoint gereed hebt, mag ik morgen dan even iets bij je afmaken aub anders krijg ik een onvoldoende') gaat over huiswerk en is in een Limburgs dialect. Het spannende van het taalgebruik op Twitter is dat het zo informeel is. Mensen schrijven op sociale media net alsof ze spreken. De huiswerktweet heeft afkortingen zoals *ff* voor 'effe' (even), *ns* voor 'eens', verdubbelingen van letters *bb* in 'aub' en het mist het onderwerp *ik* bij het werkwoord *kreijg* en het onbepaald lidwoord *een* bij het naamwoord *onvoldoende*. Dit taalgebruik laat zien wat het geschrevene kan missen (onderwerp, lidwoord) zonder (te) onbegrijpelijk te worden. Het gebruik van tekens – emoticons – is bijna een 'must'. Hoewel een lezer van een tweet de schrijver niet kan zien, kan hij of zij uit die emoticons wel opmaken hoe de schrijver zich voelt of hoe de boodschap bedoeld is. Een emoticon als :-) geeft een gezichtsuitdrukking weer. Het staat voor een lachend gezicht dat tegen de klok in 90 graden gedraaid is. Het gezicht ligt dus op zijn zijkant: als u uw hoofd naar links legt op uw schouder dan is het goed te zien. Het haakje) representeert een lachende mond, het streepje - de neus en de dubbele punt : de ogen. In de huiswerktweet representeert :p een lachend gezicht met een uitgestoken tong en x of xx staat voor een kus.

Het spannende aan Twitter is dat het een vrijplaats is om in een dialect of minderheidstaal te schrijven. Wereldwijd valt dezelfde tendens te ontdekken: talen die voornamelijk spreektaal zijn en geen stem hebben in conventionele gedrukte media zoals kranten, boeken, tijdschriften zijn wel te lezen op Twitter en Facebook. Een paar voorbeelden zijn: Welsh, Gaelic, Fries, Yucatec (een Maya taal), Euskara of Baskisch, Kreyòl ayisyen (Haïti), Setswana (Zuid-Afrika), Zwitsers Duits, Nigeriaans Pidgin, Jamaicaans Creool, Duits-Turks, Spanglish (mix tussen Spaans en Engels in VS), Iers, Sheng (mix tussen Swahili en Engels in Kenia) en een Limburgs dialect. Sociolinguïsten beweren wel eens – een beetje te optimistisch wellicht – dat sociale media, minderheidstalen voor uitsterven behoeden.

Een Kerkraadse schrijft mij waarom zij altijd dialect op sociale media gebruikt: “Kerkraads is mijn moedertaal en door ’t te schrijven (...) maak je kenbaar dat je trots bent op de gemeente of provincie waar je vandaan komt.” Dit is interessant want zo wordt in de beleving van de twitteraar zijn of haar taalgebruik op sociale media toch weer aan een specifieke plek verankerd, net als bij het dagelijkse spreken. Het voordeel om een minderheidstaal op sociale media te gebruiken, is dat niemand zich zorgen hoeft te maken over spelling en dat de tweet door de hele wereld kan zwerven. De schrijver en lezer hoeven zich niet meer op de plek te bevinden waar de taal traditioneel thuishoort. Dit is vooral belangrijk voor degenen die naar elders gemigreerd zijn, zij kunnen zich zo toch nog in hun dialect of minderheidstaal uiten en communiceren met anderen over de wereld. Zo wordt iedere minderheidstaal een globale taal.

Vragen die onderzoekers hebben, zijn waarom mensen überhaupt in het dialect of Fries twitteren en of ze dit altijd doen. Heeft dat te maken met onderwerpskeuze of met het type lezer of...? Het is immers tegendraads: dialect of minderheidstaal lijkt iets van vroeger, uit een tijdperk van de kroontjespen of het leesplankje, maar het verovert graag een plek in het moderne digitale tijdperk.

Twitteren in Limburg en Friesland

Limburgers twitteren meer in dialect dan Friezen in Fries. Dat blijkt uit het Twidentity-onderzoek¹ waarin we Friesland met Limburg vergelijken in

¹ Het Twidentity onderzoek werd gefinancierd door NWO (beurs 314-98-008). Onderzoekers in dit project waren Theo Meder (projectleider), Dolf Trieschnigg, Dong Nguyen, Lysbeth Jongbloed, Jolie van Loo en bovengenoemde auteur.

hun twittergedrag in respectievelijk het Fries en Limburgs dialect (Nguyen, Trieschnigg en Cornips 2015). In het Fries ziet zo'n tweet er zo uit: 'Ik twitterte yn it Frysk.' en in het Limburgs zo: 'Waarsjenlik noe al in Sjanje, de kroenekrane die v'r hōbbe zeen vlege op weeg daonaotoe'. Eerder is er naar het twittergedrag in het Nederlands gekeken. Nu is het zoeken naar tweets in het Limburgs veel complexer dan zoeken naar tweets in het Nederlands vanwege de variatie in de diverse dialecten en in spelling. Nederlandse tweets zijn gevonden door te zoeken op zaaiwoorden als *het* en dat werkte goed. Maar hoe dit nu te doen voor een Limburgs dialect. Welk sleutelwoord is succesvol? Een woord als *dich* maakt geen onderscheid tussen tweets in het Limburgs en in het Duits. Een zaaiwoord als *neet* 'niet' en *diech* 'jij' werkt goed. Daarmee vinden we dialecttweets uit Melick, Venlo, Roggel, Meerssen, Echt, Sint Geertruid, Geleen, Lottum en Maastricht.

Tussen januari 2013 en juli 2014 heeft de Universiteit Twente een uitvoerige databank met tweets aangelegd van twitteraars in Nederland. Het is niet echt eenvoudig om na te gaan welke tweets uit die verzameling uit Friesland en Limburg komen want tweets zijn in principe niet plaatsgebonden. Gelukkig gebruiken sommige twitteraars zelf plaatsaanduidingen. Dit is uitgezocht voor Venlo waar bijna 2.000 gebruikers aangeven dat Venlo hun twitterthuisbasis is. Het is grappig om te zien welke naam zij aan hun Venlo geven: naast het gewone V/venlo (ruim meer dan de helft), komt Venlo voor in combinatie met '(the) Netherlands', of met Blerick. Anderen benoemen Venlo eenmalig en zeer creatief als 'about 500k from Paris', 'Venlo, ja toch', 'venlo netuurlik', 'Venlocity', 'Venlo-Zuud', 'Upper E-side venlo', 'hartje Venlo', 'VenloOw', 'oet ut stedje venlo', 'earth venlo' en 'venlo wao anders?'. In deze beschrijvingen komt dus Engels voor ('about 500k from Paris'), dialect ('oet ut stedje venlo') en Nederlands ('hartje Venlo').

Een andere veel nauwkeuriger manier om te achterhalen waar iemand vandaan tuittert, is de gps-coördinaten van de twitteraar achterhalen. Die gps-coördinaten vertellen ons echter niet of de twitteraar daar woont, werkt of op bezoek of op doorreis is. In de onderzochte periode zijn er 9.843 twitteraars actief in Limburg en 11.304 in Friesland. De twitteraars in Friesland gebruiken in ruim vijf procent van hun tweets Fries en Limburgse twitteraars in bijna acht procent dialect. En Limburgers schrijven ook veel meer Engels dan dialect en ook meer dan de Friezen doen: bijna zeventien procent in Limburg laat Engels in hun tweets zien tegen ruim negen procent in Friesland. Maar het is wel knap lastig om automatisch te zoeken naar Limburgs dialect in een tweet. Limburgers schrijven veel

Nederlands en dialect door elkaar. In ‘Met 32 graden toch maar vol gaas werke aan mien stageopdracht! #onderzoek’ staan *gaas* en *mien* tussen een volledig Nederlandse zin en *werke* kan of dialect of Nederlands zijn. In de tweet ‘Toch best prima een weekendje niks! Lang geleeje!’ verraaft alleen *geleeje* dialectgebruik of zeer informeel Nederlands. In Limburg zien we een ‘powergroep’ van dialect-twitteraars: er is slechts een kleine groep die uitsluitend in dialect schrijft maar wel vaak. De meeste twitteraars in Limburg gebruiken weinig dialect tussen veel Nederlands.

Twitteren in Horst aan de Maas (Limburg)

We hebben tweets onderzocht die verzonden zijn vanuit Horst aan de Maas in het noorden van Limburg. In die tweets komt veel taalvariatie voor. De variatie die te zien is in het onbepaald lidwoord bestaat naast *een* uit: *enne*, *ennen*, *unne*, *eine*, *ein* (zie (1)):

- (1) a. Limburg bedankt! Ut waas *n* superseizoen!
- b. Melick, waat *enne* prachtige avond!
- c. Carnaval in Meterik luupt als *ennen* trein!
- d. hmmm, ik doch det die #broek van @franspollux *unne* metafoor waar...
- e. Waat *eine* geweldige hudjesbal! Top band en top artieste! En veural *ein* top dreesjpan! #genaote @devaegers
- f. Waat *een* fiest op de Parade!

Ook het bezittelijk voornaamwoord kent meer variatie dan in het Nederlands in het enkelvoud voor de eerste en tweede persoon, te weten *mien(e)* en *dien(e)*:

- (2) a. En det is waat dr *mien* inzeens aan ontbrook.
- b. grtqpie van herte mit *diene* verjeurdaag, en laot dich maar good verwinne...

Het Limburgs onderscheidt drie grammaticale geslachten (mannelijk, vrouwelijk en onzijdig) waar het Nederlands er slechts twee heeft (onzijdig *het* versus common *de*). In de streek van Venlo komt mannelijk *diene* voor, vrouwelijk *dien* en onzijdig *die*.²

2 Met dank aan Frens Bakker.

Als onderzoekers zien we vooral de Randstad als de kosmopolitische plek waar veel nieuwkomers zich vestigen, maar niets is minder waar. Horst aan de Maas in het noorden van Limburg staat bekend om grote aantallen Polen die er tijdelijk of permanent wonen en merendeels in de tuinbouw werken. Deze gemeente herbergt volgens het CBS het hoogste aandeel Poolse migranten van Limburg, ongeveer drieduizend op een bevolkingsaantal van krap veertigduizend.

Het twitteronderzoek geeft een goed beeld van de talige rijkdom in een gemeente als Horst aan de Maas. Tweets die vanuit Horst verstuurd zijn, misschien vanaf een attractiepark of benzinstation, zijn in het Duits ‘die drittsemester des #jmw studiengang der #fhkiel sind auf dem weg nach #brüssel’, in het Frans ‘dire que je suis pas loin de stro’, in het Belgisch Nederlands ‘keiplezant! @ attractiepark toverland’, in het Arabisch en Surinaams Nederlands.

Het onderzoek laat zien dat twitteraars die daar wonen en dus aan Horst gebonden zijn in het Nederlands, dialect en Engels twitteren maar niet in het Pools of in het Duits. Dat is best een verrassing gezien de vele Polen die in Horst wonen en werken. Waarom we geen Poolse tweets vinden bij degenen die in Horst wonen, kan zijn omdat de databank van de Universiteit Twente niet compleet is of omdat Polen niet twitteren. Voor de gemeente is dit hoe dan ook een belangrijk resultaat. Er is immers geen garantie dat voorlichting en informatie getwitterd door de gemeente hun Poolse inwoners ook echt weet te vinden.

Facebook: lokale identiteitsconstructie door taalculturele praktijken

Op de twaalfde van de twaalfde in 2012 trouwden meer mensen dan gebruikelijk en het was ook de datum voor iets nieuws in Limburg: de lancering van een facebookpagina *Nine-gag Op Zn Limburgs*. De welkomstwoorden van deze webpagina die vooral anderhalf jaar een bloeiend bestaan zou hebben, was: ‘Auw hoor hey kiek ozze gloednuuje pagina’. Deze internetpagina was toentertijd behoorlijk uniek want ondanks de Engels-Nederlandstalige titel schreef bijna iedereen in het dialect en was Engels nauwelijks terug te vinden. Het plezier dat gebruikers in het begin hadden, spatte ervan af. Een paar dagen na de lancering waren er al 26.547 personen die de pagina leuk vonden. Mensen haastten zich om foto’s van beroemde acteurs als Mr. Bean, Steven Seagal en Julie Andrews met bijschriften in het dialect te voorzien, zogenaamde memes, en daarmee het mondiale in het Limburgse te verankeren. Dit initiatief creëerde voor schrijvers en lezers dus hele

nieuwe mogelijkheden om hun dialect betekenisvol te gebruiken. Het dialect en het delen en beleven van humor als typisch Limburgs “Heat oas limbolandje ouch uns lol waat nemes anges begrip”(‘heeft ons Limbolandje ook eens lol om wat niemand anders begrijpt’) geeft een sterk gevoel van verbondenheid. Een dag later blijkt dat *nemes anges* (‘niemand anders’) de *Hollenjers* (‘de Hollanders’) zijn die *ozze grappe neet snappe* (‘onze grappen niet begrijpen’). De meest voorkomende tegenstelling die gebruikers op *Nine gag Op Zn Limburgs* creëren, is inderdaad die tussen Hollanders en Limburgers. De meme in Figuur 1 laat zien dat dialectspreken een belangrijk onderscheid aanbrengt vanuit het perspectief van de Limburger. De meme luidt vrij vertaald in het Nederlands: “De Hollander die over mijn dialect zeurt, die doe ik iets aan”. De meme werd geplaatst in juni 2013 door een man, ontving 1,2K reacties en werd 426 maal gedeeld.



Figuur 1: Meme Hollanders vs. Limburgers

Cornips, de Rooij en Stengs (2012) schrijven over het hoe van identiteitsconstructie die zich in het spanningsveld van tegenstellingen afspeelt die mensen zelf construeren en beleven. “Welke tegenstellingen en welke combinaties van tegenstellingen als relevant beleefd worden is voortdurend aan verandering onderhevig. Identiteit houdt onontkoombaar een even vanzelfsprekend onderscheid in met andere mensen ‘van elders’. Het maken en benadrukken van verschillen is daarmee een belangrijke voorwaarde voor gevoelens van gemeen-

schappelijkheid. In dit proces van identificatie en differentiatie gaat het zowel om ideeën die mensen over zichzelf hebben, als om ideeën die anderen over hen hebben. Stereotypering is dan ook de grondslag van de constructie van groepsidentiteiten. Mensen weten precies wat het eigene aan hun dialect is of wat hun feesten uniek voor de streek maakt.” (Cornips, de Rooij en Stengs 2012: 22). Dat zo’n facebookpagina ook van buitenaf als iets heel eigens en regionaals gezien wordt, blijkt wel uit de vrijwel onmiddellijke tegenbeweging van de Brabanders met hun *Nine-gag Op Zn Brabants* twee dagen later.

Carnaval: onderscheid met Brabanders

Tijdens het carnaval dat *vastelaovend* genoemd wordt in Limburg, hebben veel gebruikers op Facebook hun profielen van *schmink*, *pruukskes* (‘pruikjes’), *mötskes* (‘mutsjes’) en creatieve *pekskes* (‘verkleedpakjes’) voorzien.

Carnaval, een nieuwe benaming uit de zeventiende eeuw vergeleken met *Vastenavond* (*Fastnacht*) dat veel ouder is, is een feest van meedoen, een gemeenschapsfeest. De schrijver Jan van Mersbergen omschrijft het als een feest van ‘samen feestvieren, de nacht aanvallen, drinken, zingen, dansen’, of van ‘gevoelde saamhorigheid, relativering, verbroedering en loutering’. Maar carnaval is vooral ook een feest waarin het verschil met anderen zich toont, verdiept en verscherpt. *Vastelaovend* hier is niet het carnaval van daar: het is niet het armzalige en calvinistische gebeuren in Holland, de boerenkiel in Oeteldonk en het carnaval binnenshuis in gehuurde garages of boerenschuren in Oost-Nederland.



Figuur 2: Meme carnaval

Op 16 januari 2016, plaatst iemand de meme in Figuur 2 op *Nine gag Op Zn Limburgs*. Het toont de beroemde acteur Gene Wilder met bijbehorende tekst: ‘Oh, je bent een Brabander. Vertel ons alsjeblief meer over hoe goed u bent in carnaval vieren’. Deze meme wordt 73 maal gedeeld en ontvangt 44 reacties. Zo’n reactie laat zien waarom deze meme zo grappig is in de ogen van een Limburgse carnavalsvierder. De geportretteerde acteur is niet geschminkt en dat vindt degene die reageert hilarisch: “HAHAHAHA BROABANT Carnaval mot mit veul sjmienk en Limburgse muziek anders is het gein fatsoenlijk carnaval”. (‘HAHAHAHA BRABANT Carnaval moet met veel schmink en Limburgse muziek anders is het geen fatsoenlijk carnaval’).

De meme en deze reactie daarop geeft helder weer hoe Limburgse carnavalsvierders zich onderscheiden van de Brabantse: schmink is belangrijk en Limburgse muziek. Het verschil tussen hier en dáár is ook op de vierkante kilometer te maken: in Roermond is de carnavalsvereniging *d’n Uul* niet de *Veldjmuus* of de *Katers* in Maasniel (Thissen 2007).

Het zingen van de *leedjes* (carnavalsliedjes in het dialect), het spreken en schrijven van dialect verheft de beleving van carnaval als een lokaal feest, als een feest van Limburg, als iets eigens. Bij het onderscheiden van wie bij ‘wij’ en wie bij ‘zij’ hoort, gaat het zowel om ideeën die mensen over zichzelf hebben, als om ideeën die anderen over hen hebben. Bijna altijd zijn dit stereotypen of vooroordelen: bij Limburgers zit het carnaval in hun bloed en zij krijgen als vanzelf de kriebels en alleen een Limburger kan het ‘goede’ carnaval vieren. Deze stereotypen laten zien wat mensen ‘gewoon’ vinden, omdat mensen op een bepaalde plek geboren en opgegroeid zijn, en wat zij als vanzelfsprekend tot de eigen identiteit rekenen. Maar hoewel mensen er vaak zo over spreken, is identiteit geen aangeboren en vaststaand iets. Het staat in wisselwerking met de buitenwereld en omdat onze buitenwereld continu verandert, verandert ook met wie we ons al dan niet willen identificeren. Deze identificaties laten zien dat hoewel het carnaval nu als iets unieks voor Limburg ervaren wordt, dat voor de oorlog nauwelijks het geval was. De gegoede burgerij richtte met de Sociëteit Momus in Maastricht (1839) en Sociëteit Jocus in Venlo een paar jaar later wel de oudste carnavalsverenigingen op. Maar pas in de jaren vijftig van de vorige eeuw ging het georganiseerde carnaval écht leven getuige de in rap tempo opgerichte 115 nieuwe carnavalsverenigingen door heel Limburg. Bovendien: niet elke Limburger wil Vastelaovend vieren of er iets mee te maken hebben – men gaat ook op wintersport of rommelt lekker in huis – en niet elke carnavalsvierder wil zich Limburger voelen of participeert en beleeft Vastelaovend op dezelfde wijze.

Spellingsperikelen op Facebook

Wat vaak voorkomt, is dat schrijvers op deze facebookpagina reageren op elkaars dialect. Bij een meme met het opschrift “Ich hub ein augske op dich” (‘Ik heb een oogje op je’) reageert een bezoeker met “Ich han é eugske op dich zant vur bie os” (‘ik heb een oogje op je zeggen we bij ons’) waarin de variatie tussen *hub ein augske* versus *han é eugske* op de voorgrond komt. Het vele gehakketak op elkaars spelling leidt op 24 juni 2013 tot het plaatsen van de meme in Figuur 3: “Es ‘irritatie’ op de reacties die vaak te laeze zint...”. Op de meme is Robin te zien die zegt: “deze meme is niet goed geschreven” waarop Batman naar hem uithaalt en zegt “Er bestaat geen Algemeen Beschaafd Limburgs”. Deze meme ontving 344 likes en werd 22 maal gedeeld.



Figuur 3: Meme Robin en Batman

Taal is zo geschikt voor spel op sociale media omdat sprekers – en in dit geval schrijvers – van moment tot moment zichzelf steeds opnieuw kunnen neerzetten ten opzichte van anderen.

Limburgismen

Een Neerlandicus – Tommy Hopstaken – richt vervolgens op 29 maart 2013 een facebookpagina op die hij Limburgismen noemt. Deze pagina vermeldt: “Een

Limburgisme is een Limburgse uitdrukking die door Limburgers zonder blikken of blozen ‘vertaald’ wordt naar het Nederlands. Op het eerste gezicht lijkt er niks mis met deze vertaling, maar schijn bedriegt!” Bij een oude zwart-witfoto van een jongetje bijvoorbeeld staat te lezen: ‘Ik ben ontzettend *groots* op mijn nieuwe *helpen*’. *Groots* is het Limburgisme voor ‘trots’ en *helpen* voor ‘bretels’. De website toont naast idiomen als *watblief?* en *non de pie* ook woorden zoals *slips* of *sjlieps* voor ‘stropdas’, *zaal* voor ‘zadel’, *tas* voor ‘kop’, *groots* voor ‘trots’, *linksom* voor ‘binnenstebuiten’, *appelenkits* voor ‘klokhuis’, *sjiek* (kan van alles betekenen afhankelijk van de context), *kleedje* voor ‘jurk’, *extra* voor ‘expres’, *beurs* voor ‘portemonnee’, *tuitje/tuutje* voor ‘zakje’, *kwatsch* voor ‘kolder’, *zwegel* voor ‘lucifer’, *eng* voor ‘nauw’, *scheem* voor ‘schaduw’, *vaardig* voor ‘klaar’, *raaffe* voor ‘oneffenheidje’, *stevelen* voor ‘laarzen’ en *muurtjes* voor ‘wortelen’.

Er zijn ook werkwoorden te vinden zoals *inschudden* voor ‘inschenken’, *ruizelen* voor ‘uitvallen’, *foetelen* voor ‘valsspelen’, *varen* voor ‘rijden’, *kratsen* voor ‘krabben’, *opregen* voor ‘opwinden’, *aftrekken* voor ‘doortrekken’ – dit werkwoord leert iedere Limburger in vijf minuten af in de Randstad, *trampelen* voor ‘trappelende bewegingen maken’ en *pitsen* voor ‘pijn doen’. De pagina vertoont slechts twee functiewoorden: de voornaamwoordelijke bijwoorden *d’r overheer* ‘eroverheen’ en *erneven* ‘ernaast’ en het Maastrichtse voegwoord *tot* voor ‘dat’ als in ‘ik zie tot de zon komt’. Absoluut onvertaalbaar is de uitdrukking *Ik heb de knopen af*.

Conclusie

De sociale media Twitter en Facebook lenen zich uitstekend om in een minderheidstaal te schrijven. Twitteraars in Friesland en Limburg twitteren hoofdzakelijk in Nederlands, Engels en in een Limburgs dialect en Fries hoewel twitteraars in Limburg meer in Engels en dialect twitteren dan twitteraars in Friesland. Limburgers twitteren meer ‘gemixt’ tussen dialect en Nederlands. Het taalgebruik op sociale media laat veel individuele en lokale verschillen zien, bijvoorbeeld variatie in het onbepaald lidwoord *een* dat in Limburg flectie heeft voor mannelijk, vrouwelijk en onzijdig grammaticaal geslacht. Zowel de tweets als de facebookberichten tonen veel diversiteit in spelling en mensen betwisten elkaar om hun spelling. Ten slotte geven sociale media cruciale inzichten hoe regionale identiteitsconstructie tot stand komt. Limburgers onderscheiden zich van ‘Hollanders’ door het gebruik van dialect en van Brabanders in context van carnaval.

De belangrijkste onderzoeksvraag die te stellen is na dit onderzoek naar taalgebruik en identiteitsconstructies op sociale media, is hoe deze sociale media-praktijken interageren met en relateren aan gesproken taalgebruik in bijvoorbeeld Limburg.

Bibliografie

- Cornips, L. (2016). Limburg twittert. *De Limburger/Limburs Dagblad* (9 mei).
- Cornips, L. (2014). Twitteren. *De Limburger/Limburs Dagblad* (17 mei).
- Cornips, L. (2013). Vastelaovend. *De Limburger/Limburs Dagblad* (16 februari).
- Cornips, L. (2013). Limburgismen. *De Limburger/Limburs Dagblad* (14 september).
- Cornips, L., V. de Rooij & I. Stengs (2012). Carnavalesk taalgebruik en de constructie van lokale identiteiten. *Dutch Journal of Applied Linguistics* 1 (1), 15-40.
- Nguyen, D., D. Trieschnigg & L. Cornips (2015). Audience and the use of minority languages on Twitter. *Proceedings of the Ninth AAAI International Conference on Web and Social Media (ICWSM-2015)*, 666-669.
- Thissen, L. (2017, te verschijnen). *Talking in and out of place. Ethnographic reflections on language, place, and (un)belonging in Limburg, the Netherlands*. Ongepubliceerd Proefschrift Universiteit Maastricht.

Brabants dialect op Facebook en Twitter

Jos Swanenberg

Deze bijdrage behandelt het schrijven op sociale netwerksites, en dan voornamelijk in enkele specifieke gevallen op Facebook en Twitter. Zij gaat niet over grote dataverzamelingen, maar zij doet verslag van kwalitatief verkennend onderzoek naar taalgedrag op media die relatief nieuw zijn en absoluut anders dan die wij voor de intrede van het internet kenden.

Het schrijven op sociale media bestaat over het algemeen uit informeel schrijftaalgebruik. Daarom biedt dit digitale schrijven ruimte voor taalvariatie. Informeel taalgebruik op sociale media verschilt van formeel taalgebruik wat betreft de private context, nabijheid en directheid, waardoor de vorm van de berichten op sociale media op het gebied van de orthografie wat lossier is dan we traditioneel gewend zijn van schrijftaal. De orthografie van berichten op Twitter en op Facebook, althans wat betreft de gevallen die in deze bijdrage aan de orde komen, is vaak minder los dan die op andere sociale media zoals Snapchat of WhatsApp. De belangrijkste reden daarvoor is dat de berichten bedoeld zijn voor een wat groter publiek dat niet per sé bekend is aan de auteur.

Daarnaast gaat deze bijdrage over berichten die gesteld zijn in het Brabantse dialect, zoals gesproken in de provincie Noord-Brabant, in verschillende lokale vormen, bijvoorbeeld het dialect van de stad Tilburg. Dialectgebruik in Noord-Brabant staat onder druk omdat dialecten hun positie als moedertaal de laatste decennia meer en meer kwijtspelen. Dat heeft tot gevolg gehad dat er zowel horizontaal als verticaal nivellering plaatsvindt. Dat proces van taalverandering, dialectnivellering, wordt wel als volgt geanalyseerd (Britain 2009):

- > Horizontaal: lokale dialectkenmerken in woordenschat en klankenpalet worden vervangen door regionale kenmerken die allicht een groter verspreidingsgebied kennen. Een toename van dialectcontact heeft geleid tot de dominantie van kenmerken die in een groter gebied bekend zijn,

ten faveure van echt lokaal bekende kenmerken; het resultaat is een koine, een genivelleerde regionale taalvariëteit.

- > Verticaal: dialectkenmerken worden vervangen door kenmerken uit het Algemeen Nederlands, de standaardtaal in onze samenleving met een sterke positie in onderwijs, de massamedia enz.

Deze twee veranderingen leiden samen tot het verlies van lokale dialectkenmerken. Wat we in dit proces echter kunnen constateren is dat dialect zijn positie als moedertaal weliswaar meer en meer verliest, maar dat dialect niet zomaar verdwijnt. Informeel taalgebruik in gesprekken blijft een veelheid aan Brabantse kenmerken vertonen. Taalgebruik blijft dus herkenbaar Brabants, bijvoorbeeld doordat regionale kenmerken bewaard blijven (in het horizontale spectrum) of doordat er compromisvormen tussen Algemeen Nederlands en lokaal dialect ontstaan (in het verticale spectrum).

Er is dus sprake van een zekere resistentie. We weten dat allerlei bijzondere dialectwoorden verdwijnen maar niet alle domeinen zijn even kwetsbaar. Belangrijker maar veel minder opvallend is dat dialect pas echt verdwijnt wanneer het taalsysteem (met grammatica zoals woordvorming en zinsbouw) wordt opgegeven en vervangen door een algemeen Nederlands systeem. Vooral op dat gebied blijkt dialect resistent, zodat er in die koine allerlei kenmerken bewaard blijven:

- > Uitgangen die oorspronkelijk mannelijk woordgeslacht aanduiden:
dun hond, unnen hond
- > Diminutieven met uitgang -ke: *mènneke*
- > De typische werkwoordvormen van de 2^e persoon: *hedde gij, doede gullie*
- > In de woordenschat bijvoorbeeld het groetwoord *houdoe*

Die taalkenmerken functioneren dan ook als sjibbolets; zij zorgen er voor dat het taalgebruik herkenbaar Brabants blijft. Dit taalgebruik noemen we nog steeds dialect, maar het is iets anders dan de traditionele dialecten met hun lokale dialecteigenschappen. We hebben inmiddels behoorlijk wat aanwijzingen dat de koine een consolidering laat zien (Wilting et al. 2014), dat wil zeggen dat we niet een stadium bereiken waarin alle Brabanders hetzelfde Nederlands spreken, laat staan dat alle Nederlanders dat gaan doen.

Be like

Uitgangspunt was voor ons: hoe ziet dialectgebruik op sites als Twitter en Facebook eruit, en komt dat overeen met mondeling taalgebruik? Op de eerste plaats hebben we gekeken naar lokale of regionale Be like-pagina's.

Op 3 januari 2014 werd *Tattas be like* gelanceerd, een facebookpagina die werd afgeleid van de Amerikaanse facebookpagina *Niggas be like*. Na drie dagen had *Tattas be like* al 60.000 likes. Het thema van de pagina bestaat uit het weergeven en bespotten van de Nederlandse identiteit met behulp van stereotypen, zoals de Nederlandse gierigheid en zuinigheid. Kort daarna verschenen er allerlei hierop geïnspireerde pagina's, die de identiteit en cultuur van bepaalde groepen in onze samenleving willen uitdrukken en bespotten. Zo volgde op 5 januari 2014 bijvoorbeeld *Surinamers be like*, op 7 januari *Vrouwen be like* en op 8 januari volgden *Studenten be like* en *Brabanders be like*.

Met die laatste spin-off wordt het regionale aspect van groepsculturen betrokken in dit verschijnsel. Eind januari 2014 constateerde Omroep Brabant al elf Brabantse facebookpagina's: *Brabanders be like*, *Brabo's be like*, *Eindhovenaren be like*, *Tilburgers be like*, *Bosschenaren be like*, *Bredanaars be like*, *Helmonders be like*, *Ossenaren be like*, *Bergenenaren be like*, *Zevenbergen be like* en *Roosendaalers be like* (Van Doorn 2014).

Op deze *Be like*-pagina's worden foto's getoond waarin korte teksten zijn geplaatst. Men noemt ze internetmemes. Het onderwerp verbeeldt en verwoordt iets dat typisch is voor een bepaalde plaats of streek, de pagina's willen het publiek dus iets vertellen over kenmerken van regionale identiteit in Noord-Brabant. Beeld en tekst vormen samen communicatieve uitingen waarin op humoristische, parodiërende wijze bepaalde kenmerken sterk uitvergroot worden. Om dit effect te bereiken wordt gebruikgemaakt van herkenbare beelden (een bekend gebouw, bekende persoon, een streekgerecht), clichématige teksten (vaak in het dialect) en stereotiepe eigenschappen (Brabanders zijn Bourgondisch, dus wordt er veel en goed gedronken en gegeten).

De moerasdraak in Figuur 1 is het symbool voor de stad 's-Hertogenbosch en de tekst 'menne schoonma is ok een draak' is in het dialect gesteld, maar merk op dat er een fout tegen de grammatica in zit: *schoonma* is niet een mannelijk woord dus zou *menne* niet de uitgang die het mannelijk woordgeslacht markeert, moeten hebben. Men zou *menne schoonma* verwachten. Anderzijds is juist dat bezit-

telijk voornaamwoord vanwege die uitgang het woord dat de tekst bij uitstek niet-nederlands maakt; het laat zien dat deze zin in het dialect is geschreven. Het is dus een dialectvorm, die in het Bosch en in het Brabants voorkomt, maar in deze tekst past de vorm eigenlijk niet. Zulke verkeerd gebruikte, overdreven dialectvormen worden hyperdialect genoemd.



Figuur 1: Meme moerasdraak

Zodoende worden vormen als *unne*, *den* en hier *menne* als typisch Brabants dialectkenmerk gezien, als sjibbolet herkend en vervolgens hypercorrect gebruikt. Zo worden hyperdialect en ook bijzondere spellingswijzen (*quekske* in plaats van *kuukske* voor 'koekje') gebruikt om een overdreven, stereotiepe Brabantse identiteit neer te zetten, wat precies past bij de intentie van deze facebookpagina's (Visser et al. 2015). In de volgende paragraaf gaan we in op dialectgebruik op een ander sociale netwerk, Twitter.

De Braboneger

Braboneger is het pseudoniem van Steven Brunswijk, een stand-upcomedian die in eerste instantie bekend werd via clips op YouTube en vergelijkbare kanalen. Hij is geboren in Suriname, maar woont sinds zijn derde in Tilburg in Noord-Brabant. Sinds november 2016 is de Braboneger ook op het nationale televisiekanaal NPO 3 te zien met zijn programma *De Braboneger verkaast!?*. In zijn eerste clip op YouTube vertelt hij onder meer hoe Surinaamse mensen maar klagen over de economische crisis en dat ze daarom hun rekeningen niet

langer betalen, maar dat ze eens zouden moeten denken aan de arme mensen in Afrika, en hoe ver zij moeten lopen om water te halen: dus hou op met klagen, ga werken, vergeet je uitkering, en betaal gewoon je rekeningen. In feite verwoordt Braboneger in plat Tilburgs de mening van de ‘gemiddelde blanke man in de straat’, een mening die je misschien niet zo vaak hoort omdat die als discriminerend wordt beschouwd, maar als een Surinamer dat zelf zegt, geeft het een humoristisch effect. De eerste live show van deze artiest heette dan ook ‘Vooroordelen’. Hij speelt dus met zwart en wit, een man met het uiterlijk van een Surinamer, die denkt en praat als een autochtone Tilburger ‘van de straat’. Hij sluit deze sketch af met “overigens spreek ik ook keurig ABN”. Braboneger spreekt dialect in zijn videoclips en ook op het podium. En dat doet hij ook heel consequent op Twitter.



Figuur 2: Tweet Braboneger over de Verenigde Naties

In de tweet in Figuur 2 zien we, net als bij de Be like-pagina's hierboven, de markering van het woordgeslacht als afwijkend dialectkenmerk. Het woord *onderzoek* is immers onzijdig maar toch krijgt het onbepaald lidwoord *unne* een uitgang, die suggereert dat het woord mannelijk is. Eigenlijk zou men moeten schrijven *un onderzoek*. Die vorm zou er evenwel meer Nederlands uitzien dan de vorm die nu is geschreven; nu ziet het er meer Brabants uit. (Een ander voorbeeld hiervan is *den televisie*.)

Ook in Figuur 3 zien we een vorm, *kopske*, die er Brabants uitziet met de uitgang *-ske*, maar die niet overeenstemt met de traditionele dialectgrammatica, op grond waarvan we *kopke* of *köpke* verwachten. (Andere voorbeelden zijn *mopske* voor 'schatje', zoals men ook wel *moppie* zegt, en *fiske* waar *fisje* voor 'feestje' meer gebruikelijk is.) De uitgang *-ske* hoort alleen bij woorden eindigend op velleire medeklinkers, zoals bij *bak-bèkske*. Ook hier wordt een opvallend dialectkenmerk, de uitgang *-ske* voor verkleinwoorden, weer gebruikt waar we dat niet verwachten.



Figuur 3: Tweet Braboneger over zonnestudio

De tweets van de Braboneger bevatten dus dialectkenmerken op plaatsen waar we ze niet verwachten, alsmede schrijfwijzen die een uitspraak tussen dialect en Nederlands in, compromisvormen, suggereren en daarnaast is een aantal woorden gewoon op Nederlandse wijze gespeld.

Is dit dialectgebruik deel van de act die deze man opvoert? Dit vindt hij er zelf van: “Ik kwats maar wat. Vooral met *aaawe mêense* kan ik goed Brabants lullen, veel jongeren kunnen dat zelf vaak al niet meer. Ik praat alleen ABN als het écht niet anders kan. Als Braboneger praat ik in elk geval niet anders dan normaal, ik hoef me niet te forceren. Ik zet het soms alleen iets dikker aan” (Brabants Dagblad, 14 maart 2014).

Conclusie

We zien dat teksten van internetmemes op Facebook en berichten op Twitter een dialectgebruik laten zien dat ver af staat van bijvoorbeeld het authentieke Tilburgs. Dat authentieke Tilburgs is behoorlijk goed bekend. Het is veruit het best beschreven dialect van de provincie Noord-Brabant met uitgebreide dialectwoordenboeken, met een grammatica, en zelfs een proefschrift over de fonologie. Daarnaast is het Groot Tilburgs Diktee een populair jaarlijks evenement waar honderden mensen aan meedoen, je kunt schrijfcursussen volgen en er zijn diverse educatieve projecten voor basisscholen. Dat dialect is een totaal ander dialect dan de voorbeelden uit deze bijdrage laten zien. Er gelden spelingsvoorschriften en er zijn normen voor dat authentieke Tilburgse dialect: men weet hoe het hoort. Maar het lijkt ook wel te worden beschouwd als een dode taal, als een stabiel onveranderlijk systeem; men houdt nauwelijks rekening met taalverandering.

Je kunt je natuurlijk afvragen hoe authentiek het Tilburgs van Braboneger is. Maar wat doet dat er eigenlijk toe? Zijn doel is entertainment; als het Tilburgs klinkt, of het er Tilburgs uitziet op Twitter, dan past zijn taalgebruik al bij zijn act. Hij spreekt nauwelijks anders dan dat hij schrijft. Het Tilburgs van Braboneger is onderdeel van zijn geacteerde Tilburgse identiteit, maar dat lijkt niet veel anders te zijn dan hoe hij zich in het dagelijks leven gedraagt. Braboneger zegt daar zelf over:



Figuur 4: Tweet Braboneger over taalgebruik

Het imago van Brabant of 's-Hertogenbosch zoals dat wordt neergezet in de memes van *Be like*-pagina's op Facebook is heel stereotiep. De manier waarop dialect wordt ingezet past bij dat stereotiepe imago; het Brabants wordt als duidelijk anders dan het algemeen Nederlands ervaren, wordt zodoende ingezet voor humoristische doeleinden en in de overdrijving van de stereotypering worden dialectkenmerken overdreven vaak gebruikt (hyperdialect). Dialect maakt deel uit van het stereotype. Op deze wijze worden social media ingezet voor uitingen van regionale of lokale taal en de constructie van een regionaal of lokaal imago. Maar het laat ook zien hoe het nieuwe Brabants er 'in het echt' uitziet. De kenmerken die ik hier behandeld heb, vinden we ook in de spreektaal van jonge Brabanders op middelbare scholen, en zo werden ze trouwens ook al opgetekend door Cor Hoppenbrouwers bij zijn onderzoek naar regiolectisering eind jaren 80. Deze hyperdialectvormen zijn typisch voor dit proces van taalverandering van lokaal dialect naar regionaal koine.

De compromisvormen en vooral het hyperdialectgebruik blijken hardnekkig, lijken wel toe te nemen, laten zien dat sommige kenmerken resistent zijn, vooralsnog, maar daarbij wel van functie kunnen veranderen. De sociolinguïstische functie lijkt mij de uiting van regionale identiteit; je laat ermee zien dat je een Brabander bent.

Bibliografie

- Britain, D. (2009). One foot in the grave?: Dialect death, dialect contact and dialect birth in England. *International Journal of the Sociology of Language* 196/197, 121-155.
- Hoppenbrouwers, C. (1990). *Het regiolect. Van dialect tot algemeen Nederlands*. Muiderberg: Coutinho.
- Van Doorn, B. (2014). Be Like nieuwste hype op Facebook: 'Brabanders Be Like... Dè Minde Nie!'. *Omroep Brabant* (28 januari). Geraadpleegd op 23 mei 2014 via <http://www.omroepbrabant.nl/?news/205901592/Be+-Like+nieuwste+hype+op+Facebook+Brabanders+Be+Like...+D%C3%A8+Minde+Nie!.aspx>
- Visser, P., J. Nortier & J. Swanenberg (2015). Dialect als stereotypering. Regionale identiteit in Noord-Brabant verbeeld en verwoord op Facebook. In M. Curcic, S. Bacchini, B. van den Bogaerde & M. Boogaard (red). *Papers of the Anéla 2015 Applied Linguistics Conference*. Delft: Eburon, 111-124.
- Wiling, M., R. van Hout & J. Swanenberg (2014). Regiolect verankerd. Een survey-onderzoek naar dialectgebruik van Eindhovense jongeren. *Taal en Tongval* 66 (2), 143-171.

Het gebruik van Friese varianten op sociale media

Lysbeth Jongbloed-Faber

Sinds 2013 doe ik onderzoek naar het gebruik van Fries op sociale media. In onze dagelijkse communicatie nemen sociale media een steeds belangrijkere plek in, en aangezien Fries voornamelijk gesproken wordt en in geringe mate geschreven, is het interessant om te onderzoeken hoe het Fries wordt gebruikt op deze nieuwe media. Sociale media verbinden mensen, onafhankelijk van tijd en plaats. Aan de ene kant biedt dit de mogelijkheid aan mensen die vertrokken zijn uit Friesland om toch in hun moedertaal te kunnen blijven communiceren; aan de andere kant is er op sociale media vaak een verscheidenheid aan talen aanwezig en eerder onderzoek naar het gebruik van minderheids- en migrantentalen toont aan dat er vaak druk wordt uitgeoefend om een dominante taal te gebruiken (Androutsopoulos 2014). Onderzoek in Wales naar het gebruik van Engels en Welsh laat zien dat het taalgebruik op sociale media vaak het offline gebruik weerspiegelt: in een-op-eenberichten wordt vaak het Welsh gebruikt terwijl als iemand grotere groepen wil aanspreken vaker het Engels wordt gehanteerd (Cunliffe, Morris & Prys 2013).

Hoe gevarieerder de taaldiversiteit van iemands publiek, hoe complexer het wordt om het publiek goed aan te spreken. Marwick & Boyd (2011) spreken over de 'context collapse': op sociale media komen groepen van mensen samen die elkaar in de offline wereld nooit tegelijkertijd zouden ontmoeten. Bovendien is er op sociale media sprake van onzekerheid over het publiek dat bereikt wordt: het is van tevoren niet duidelijk wie een bericht gaat lezen of erop gaat reageren. Bij het samenstellen van het bericht maakt de zender dan ook gebruik van het concept 'imagined audience', het geprojecteerde publiek dat iemand denkt te bereiken (Marwick & Boyd 2011).

Regionale taalvarianten worden niet langer beschouwd als een automatische consequentie van waar iemand is geboren of opgegroeid, maar eerder als een

bron voor sociale actie (Johnstone 2004). Door het gebruik van specifieke taalvarianten of stijl kunnen sociale identiteiten worden geconstrueerd, soms geassocieerd met een bepaalde plek of regio (Johnstone 2010). Het gebruik van Friese varianten op sociale media kan dus worden ingezet om (een) sociale identiteit(en) te construeren.

In de provincie Friesland heeft 55% van de 650.000 inwoners het Fries als eerste taal thuis geleerd. Fries is vooral een gesproken taal: 67% van alle inwoners spreekt goed Fries tegenover slechts 15% die het goed schrijft (Provincie Fryslân 2015). Het Fries wordt met name in informele domeinen en op het platteland gebruikt, terwijl in de steden en formele domeinen het Nederlands dominant is. Het gesproken Fries wordt gekenmerkt door veel dialectische variatie waarbij drie Friese dialectgroepen te onderscheiden zijn: het *Klaaifrysk* (Kleifries), gesproken op de kleigronden in het noordwesten van de provincie, het *Wâldfrysk* (Woudfries) dat in de Friese Wouden in het oosten van de provincie wordt gesproken en het *Súdwesthoeksk* (Zuidwestfries) dat in Zuidwest-Friesland wordt gesproken (Hoff 1933, in Stefan, Klinkenberg & Versloot 2015).

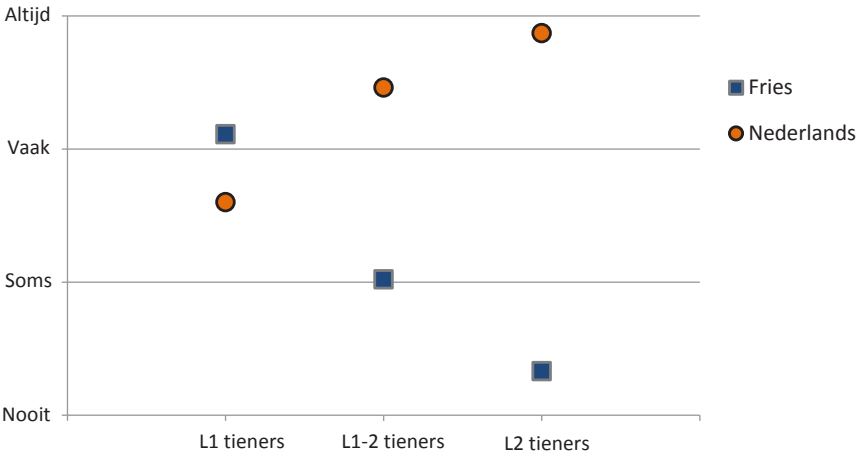
Het taalgebruik van tieners op sociale media

In 2013 en 2014 heb ik kwantitatief onderzoek uitgevoerd naar het taalgebruik van tieners in Friesland. 22 scholen in het voortgezet onderwijs en op het MBO hebben meegewerkt aan het onderzoek, waarvan 10 scholen in de steden en 12 scholen op het platteland. In totaal vulden 2.267 jongeren van 14 tot en met 18 jaar de vragenlijst in. De thuistaal van de tieners was voor 45% uitsluitend Fries (L1 tieners), voor 36% uitsluitend Nederlands (L2 tieners) en voor 10% gemengd Fries-Nederlands (L1-2 tieners). Daarnaast was 6% van de jongeren (gedeeltelijk) in een andere taal dan het Nederlands of Fries opgevoed. Voor alle groepen betrof het gerapporteerd taalgebruik.

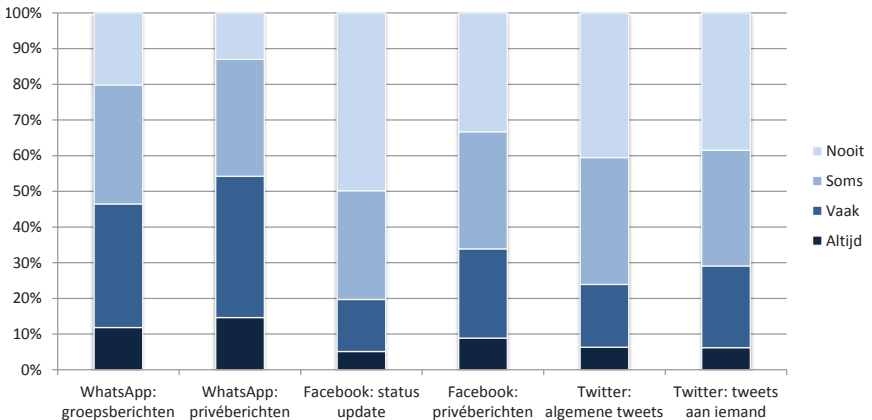
In Figuur 1 staat het gebruik van Nederlands en Fries afgebeeld in gesprekken met vrienden, uitgesplitst naar thuistaal. Alleen tieners die thuis uitsluitend Fries spreken, spreken vaker Fries dan Nederlands met hun vrienden. Terwijl 45% van de tieners vaak of altijd Fries spreekt met vrienden, ligt het aandeel Fries op sociale media een stuk lager. Op WhatsApp wordt het Fries het vaakst gebruikt: 25% gebruikt vaak of altijd Fries in groepsapps; in persoonlijke apps ligt dit aandeel op 29%. Op Facebook gebruikt 10% vaak of altijd Fries in statusupdates en 18% in chatberichten. Op Twitter wordt door 13% van de jongeren vaak of

altijd Fries gebruikt in algemene tweets en in tweets aan een specifiek iemand (beginnend met @ of in een privébericht) ligt dit aandeel op 15%.

In Figuur 2 staat het gebruik van Fries door de tieners die thuis uitsluitend Fries spreken afgebeeld. De variatie is groot, zowel per sociaal medium als ook tussen de verschillende sociale media. Gebruikt 55% van deze tieners vaak of altijd Fries in privéberichten op WhatsApp, en slechts 13% nooit, op Facebook schrijft slechts 20% zijn statusupdates in het Fries tegenover 50% nooit.



Figuur 1: Gebruik van Nederlands en Fries in gesprekken met vrienden door L1, L1-2 en L2 tieners



Figuur 2: Het gebruik van Fries op sociale media door L1 tieners

Valt het gebruik van Fries op sociale media door de tieners te voorspellen? De volgende vier factoren verklaren 56% van de variatie ($R^2=0,56$) in het gebruik:

- > Fries spreken met vrienden
- > Eigen attitude t.o.v. Fries
- > Schrijfvaardigheid
- > Attitude van klasgenoten t.o.v. Fries

De thuistaal bleek geen voorspellende waarde te hebben. Voor meer informatie over het onderzoek naar het taalgebruik van Friese tieners op sociale media verwijs ik u naar Jongbloed-Faber, Van de Velde, Van der Meer & Klinkenberg (2016).

Het gebruik van Friese varianten op sociale media

Het gesproken Fries wordt gekenmerkt door grote dialectische variatie. ‘Zaterdag’ is een van de woorden waarbij de variatie binnen Friesland het grootste is. *Sneon* [snö.ən] is de meest gebruikte voorkeursvariant. Deze variant wordt gebruikt door de Friese media, officiële instanties en de bevolking die is opgegroeid op de kleigrond en in Zuidwest-Friesland. *Saterdei* [sa:trdi] is de schrijftaalvariant uit de Friese Wouden en *saterje* [sa:trjə] de fonetische variant van *saterdei*. Door het analyseren van een aantal tweets wil ik laten zien hoe taalgebruik op het internet kan worden gelinkt aan (delen van) Friesland en hoe een lokale sociale identiteit kan worden geconstrueerd door taalgebruik op Twitter.



Figuur 3: Tweet Greidhoekfestival

De tweet in Figuur 3 gaat over het *Greidhoek' Festival*. De *Greidhoeke* is een regio op de Friese kleigrond in midden-Friesland die bekend staat om het onein-

dige weidelandschap met terpdorpen. De tweet vraagt mensen ‘met wie wil jij komende zaterdag naar het Greidhoek’ Festival?’ en roept de volgers op de promo te delen zodat men twee kaarten kan winnen voor het festival. Dat in deze tweet over een evenement op de kleigrond *sneon* wordt gebruikt is erg logisch. *Saterje* of *saterdei* zou hier volledig ‘out of place’ zijn geweest.

De polsstokhoogspringclub uit Burgum twittert dat de Koningsspelen op zaterdag 16 juli zullen plaatsvinden (Figuur 4). In de tweet is *saterdei* geselecteerd, de schrijftaalvariant van de Woudfriese zaterdag. Ook de rest van de tweet is bijna volledig in Standaardfries geschreven. Door het gebruik van *saterdei* wordt duidelijk dat de polsstokhoogspringwedstrijd in de Friese Wouden zal plaatsvinden, ook als de plaats Burgum niet zou zijn genoemd.



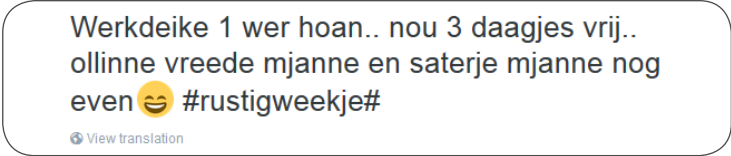
Figuur 4: Tweet Ljeppersklup

De Hûnekop is een Friese band die Friestalige ‘arbeidersrock’ speelt. De roots van deze band liggen in de Friese Wouden. In de tweet in Figuur 5 construeert de band een op en top Woudfriese identiteit door het gebruik van *saterje*. Tevens wordt er op een onconventionele, coole manier met taal omgegaan. Niet alleen door het codeswitchen van Fries naar Nederlands (*Nix loos!*), maar ook door het gebruik van de ‘x’ in die Nederlandse expressie en het gebruik van *eem* (‘even’). De band construeert hiermee een relaxte, open, Woudfriese identiteit.



Figuur 5: Tweet de Hûnekop

De tweet in Figuur 6 is een voorbeeld van fonetisch Woudfries gecombineerd met Nederlands. De jongedame codeswitcht twee keer binnen 140 tekens. Het fonetische Fries is kenmerkend voor de Friese Wouden: *vreede mjanne* ('vrijdagochtend') en *saterje mjanne* ('zaterdagochtend') wordt alleen diep in de Friese Wouden zo uitgesproken. Het afwisselen van het fonetische Woudfries met Nederlands laat zien hoe tweetalig de meeste Friezen zijn, maar kan voor de twitteraar ook een manier zijn om haar Nederlandstalige contacten aan te spreken.



Figuur 6: Tweet Woudfries

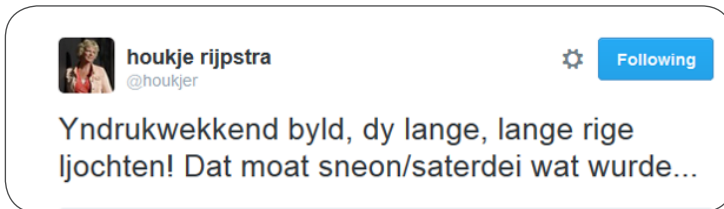
Het damesteam van voetbalvereniging Drachtster Boys twittert in het Nederlands over een succesvolle wedstrijd in de regen (Figuur 7). Echter, aan het eind van de tweet wordt de Woudfriese identiteit van het team geconstrueerd door het gebruik van de fonetische Woudfriese variant *oant saterje* ('tot zaterdag'). Veel Friezen hebben moeite met het schrijven en lezen in hun moedertaal, maar deze uitdrukking begrijpt iedereen. Hiermee wordt de tweet toch nog lokaal gemaakt.



Figuur 7: Tweet Drachtster Boys

De laatste voorbeeldtweet (Figuur 8) is afkomstig van een Friese politica die wethouder is in de gemeente Tytsjerksteradiel, geografisch gepositioneerd in de Friese Wouden. De politica spreekt en schrijft normaal gesproken Standaardfries. Bij het beschrijven van het evenement dat aanstaande zaterdag zal plaats-

vinden in de Friese Wouden lijkt de politica geen keuze te kunnen maken tussen de officiële standaard schrijftaalvariant en de schrijftaalvariant uit de Friese Wouden en selecteert beide varianten in haar tweet. Hiermee laat ze zien dat ze een competente schrijver van het Fries is, maar construeert ze tegelijkertijd een Woudfriese identiteit.



Figuur 8: Tweet Houkje Rijpstra

Het gebruik van *sneon*, *saterdei* of *saterje* geeft al veel informatie over de lokale herkomst van een twitteraar prijs. Soms zal een twitteraar bewust met deze keuze omgaan, in andere gevallen niet. In beide gevallen wordt met de keuze voor een van de varianten een regionale identiteit geconstrueerd die te linken is aan (een specifieke plek in) Friesland. Hoewel het internet een globale plek is, laat deze casestudie zien hoe talige praktijken op Twitter lokale identiteit kunnen construeren.

Effect van taalpromotiecampagnes

Ten slotte wil ik ook nog onderzoek delen dat ik heb uitgevoerd naar het effect van taalpromotiecampagnes in Friesland, en in het bijzonder het effect van de *Fryske Twitterdei*, de dag waarop iedereen wordt aangemoedigd het Fries op Twitter te gebruiken, vergezeld door #Frysk. Op 18 april 2013 werd de *Fryske Twitterdei* voor de tweede maal georganiseerd door de taalpromotiecampagne PraatMarFrysk. In totaal werden bijna 10.000 tweets verstuurd met #Frysk. Om het effect van de twitterdag te meten heb ik ruim 6.000 tweets bestudeerd van 50 tieners die op 18 april 2013 in hun tweets minimaal één keer #Frysk hebben gebruikt: per tiener heb ik de 50 tweets voor de twitterdag, alle tweets op de twitterdag en 50 tweets na de twitterdag geanalyseerd. Zie Tabel 1 voor een overzicht van de gebruikte talen in de tweets.

Taal	Voor de twitterdag	Op de twitterdag	Na de twitterdag
Meest Fries	12%	53%	13%
Meest Nederlands	65%	29%	63%
Meest Engels	8%	4%	8%
Anders	15%	14%	16%

Tabel 1: Taalgebruik voor, op en na de Fryske Twitterdei in de tweets (n=6.019)

Voor de Fryske Twitterdei werd door de bestudeerde tieners het Nederlands verreweg het vaakst gebruikt: in 65% van de tweets was het Nederlands de belangrijkste taal en in slechts 12% het Fries. Op de twitterdag zelf wordt het Fries de belangrijkste taal in de tweets: het gebruik van het Fries verviervoudigt tot ruim de helft van alle tweets (53%) terwijl het aandeel Nederlandse tweets daalt naar 29%. Direct na 18 april normaliseert het taalgebruik weer; er is zelfs geen kort nawee-effect te ontdekken.

In de introductie omschreef ik al dat onderzoek van Cunliffe, Morris en Prys (2013) aantoonde dat het gebruik van Welsh op Facebook substantiëler is in privéberichten dan in statusupdates. Door het taalgebruik per type tweet te analyseren, kan een vergelijking worden gemaakt met die resultaten. Zie Tabel 2 voor een overzicht van het aandeel Friese tweets per type tweet.

Type bericht	Voor de twitterdag	Op de twitterdag	Na de twitterdag
Algemene tweets (n=3.019)	8%	53%	9%
Geadresseerde tweets (n=1.834)	25%	47%	26%
Retweets (n=1.166)	4%	58%	4%

Tabel 2: Aandeel Friese tweets voor, op en na de Fryske Twitterdei, per type tweet (n=6,019)

Voor en na de twitterdag is het aandeel Friese tweets in de geadresseerde tweets (beginnend met @) veel hoger dan de algemene, publieke tweets en de retweets. Dit is in lijn met de bevindingen van Cunliffe, Morris en Prys (2013). Opvallend is dat, hoewel het aandeel Friese tweets op de twitterdag bij alle soorten tweets stijgt, het aandeel Friese tweets op de twitterdag zelf het hoogst is in retweets en juist het laagst in de geadresseerde tweets. Een mogelijke verklaring is dat de geadresseerde berichten zijn geschreven in de taal die het meest natuurlijk

aanvoelt voor zowel de zender als de ontvanger en dat een extern fenomeen zoals een georganiseerde *Fryske Twitterdei* minder effect heeft op de geselecteerde taal in dit type berichten. Bovendien lijkt het alsof de twitterdag het voor mensen rechtvaardigt om meer algemene tweets en retweets in het Fries op Twitter te versturen, ook al zal een groot gedeelte van het publiek de berichten niet begrijpen. Voor meer informatie over het effect van de *Fryske Twitterdei* verwijst ik naar Jongbloed-Faber (2014).

Conclusies

Kan het gebruik van Fries op sociale media worden voorspeld? Uit het kwantitatieve onderzoek onder Friese tieners blijkt dat iemands offline en online vrienden uitermate belangrijk zijn bij het gebruik van Fries op sociale media. Ook schrijfvaardigheid en attitudes tegenover het Fries spelen een rol bij iemands taalkeuze op sociale media. Het Fries wordt, net als andere minderheidstalen, vaker gebruikt in een-op-eenberichten dan wanneer iemand een groter publiek adresseert. Een analyse van taalgebruik op Twitter laat zien dat het gebruik van Friese varianten lokale, sociale identiteit(en) kan construeren. Taalpromotiecampagnes zoals het organiseren van een twitterdag kunnen het gebruik van Fries op sociale media stimuleren, maar het effect is van korte duur.

Bibliografie

- Androutopoulos, J. (2014). Linguaging when contexts collapse: Audience design in social networking. *Discourse, Context and Media* 4-5, 62-73.
- Cunliffe, D., D. Morris, & C. Prys (2013). Investigating the differential use of Welsh in young speakers' social networks: A comparison of communication in face-to-face settings in electronic texts and on social networking sites. In E.H.G. Jones & E. Uribe-Jongbloed (red.). *Social media and minority languages: Convergence and the creative industries*. Bristol: Multilingual Matters, 75-86.
- Johnstone, B. (2004). Place, globalization, and linguistic variation. In C. Fought (red.). *Sociolinguistic variation: Critical reflections*. New York: Oxford University Press, 65-83.
- Johnstone, B. (2010). Indexing the local. In N. Coupland (red.). *The handbook of language and globalization*. Malden, MA: Wiley-Blackwell, 386-405.
- Jongbloed-Faber, L. (2014). Social media: a treasure trove for minority language research. In K. Woodfield (red.), *Social Media in Social Research: Blogs*

- on Blurring the Boundaries*. London: Natcen Social Research, 189-194.
- Jongbloed-Faber, L., H. Van de Velde, C. Van der Meer & E.L. Klinkenberg (2016). Language Use of Frisian Bilingual Teenagers on Social Media. *Treballs de Sociolingüística Catalana* 26, 27-54.
- Marwick, A. & Boyd, D. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society* 13 (1), 114-133.
- Provincie Fryslân (2015). *De Fryske taalatlas 2015. Fryske taal yn byld*. Leeuwarden: Provincie Fryslân.
- Stefan, M. H., E.L. Klinkenberg & A.P. Versloot (2015). Frisian sociological language survey goes linguistic: Introduction to a new research component. In A. J. Brand, E. Hoekstra, J. Spoelstra, & H. Van de Velde (red.). *Philologia Frisica Anno 2014. Lêzings fan it tweintichste Frysk Filologekongres fan de Fryske Akademy op 10. 11 en 12 desimber 2014, vol. 1091*. Ljouwert: Fryske Akademy & Afûk, 240-257.

#dialect

Culturele modellen en de perceptie van taalvariëteiten op Twitter

Karliën Franco en Laura Rosseel¹

Mensen denken over de werkelijkheid in de vorm van culturele modellen (Holland & Quinn 1987). Deze modellen, die ook wel *folk theories* worden genoemd (Dirven et al. 2007), zijn gedeelde denkpatronen die ons helpen orde te scheppen in de sociale realiteit. Dat inzicht, dat oorspronkelijk uit de culturele antropologie komt, heeft ook linguïsten geïnspireerd om onderzoek te doen naar culturele modellen van taal en taalvariatie.² Zo gebruikt bijvoorbeeld Geeraerts (2003; maar zie o.a. ook Polzenhagen & Dirven (2008) en Berthele (2011) voor recenter werk) de theorie rond culturele modellen om vanuit de Cognitieve Sociolinguïstiek onderzoek te doen naar de overtuigingen die mensen hebben over taal en taalvariatie. Geeraerts (2003) stelt twee types denkpatronen rond taal voor: een rationalistisch en een romantisch model. In een rationalistisch model wordt taal in grote mate als een communicatiemiddel beschouwd, terwijl een romantisch model taal eerder ziet als een medium om identiteit uit te drukken.

Die denkpatronen over taal brengen ook bepaalde percepties van taalvariatie met zich mee. Een rationalistisch model stelt standaardisering bijzonder positief voor omdat een standaardtaal gezien wordt als een uniform middel dat mensen uit verschillende sociale groepen en geografische locaties toelaat over alle mogelijke onderwerpen met elkaar te communiceren. Binnen de 18^{de}-eeuwse context van de Verlichting, waar dit denkpatroon zijn oorsprong vindt, werd standaardisering

1 Karliën Franco: QLVL, KU Leuven; Laura Rosseel: QLVL, KU Leuven en Fonds voor Wetenschappelijk Onderzoek Vlaanderen.

2 Merk op dat ook in de taalkunde reeds lang wordt nagedacht over de relatie tussen taal, denken en cultuur. Zo vinden we bijvoorbeeld al in de negentiende eeuw bij Humboldt ideeën die de relatie tussen die drie componenten benadrukken (Dirven et al. 2007).

daarom beschouwd als de manier bij uitstek om sociale en politieke emancipatie te bereiken (Geeraerts 2003, 5-6). Dialecten worden daar veel negatiever gepercipieerd, omdat hun geografische en sociale beperktheid die universele communicatie en emancipatie in de weg staat (Geeraerts 2003, 5-6). Kort samengevat staat een rationalistisch cultureel model van taal negatief tegenover taalvariatie, omdat het hinderlijk is voor de communicatie.

Een romantisch model, daarentegen, ziet taal voornamelijk als een expressiemiddel. In dat denkpatroon wordt standaardisering als een negatief proces gezien, omdat de standaardtaal voornamelijk gebruikt wordt in domeinen als wetenschap, cultuur en politiek en bijgevolg vooral beheerst wordt door een beperkte sociale elite (Geeraerts 2003, 11). Het doel van emancipatie wordt in de praktijk dus niet bereikt en standaardisering leidt eerder tot exclusie dan inclusie van bepaalde sociale of geografische groepen (Geeraerts 2003, 11-12). Taalvariatie, bijvoorbeeld in de vorm van lokale dialecten, wordt in dit cultureel model als de ideale manier beschouwd om emoties te uiten en zo gestalte te geven aan de eigen identiteit in tegenstelling tot de uniforme en elitaire standaardtaal. Taalvariatie wordt dus niet als hinderlijk ervaren, zoals in het rationalistische model dat heldere communicatie voorop zet, maar als een vorm van individuele expressie.

In zijn paper uit 2003 kiest Geeraerts voor een top-down aanpak waarbij hij vanuit de theorie vertrekt en die vervolgens illustreert met een aantal voorbeelden. In deze studie opteren wij voor de omgekeerde werkwijze: we werken bottom-up en vertrekken vanuit de data om na te gaan óf en welke aanwijzingen we vinden voor rationalistische, romantische of andere culturele modellen van taal. Daarenboven beperken we het bereik van deze studie tot culturele modellen van het concept 'dialect'. We kozen dialect als het eerste concept om op deze manier te onderzoeken, omdat we ervan overtuigd zijn dat het voor de meeste Vlamingen en Nederlanders een cognitief reëel en saillant concept is binnen de context van taalvariatie.

De data die we gebruiken in deze studie zijn afkomstig van het sociale medium Twitter, een berichtenservice waarmee gebruikers korte berichten (140 tekens) de wereld kunnen insturen, al dan niet voorzien van hashtags die de berichten of tweets volgens thema categoriseren. Twitterberichten vormen een interessante gegevensbron om culturele modellen te bestuderen, aangezien Twitter het medium bij uitstek is waarop gebruikers hun mening delen, met elkaar in discussie treden en bepaalde (publieke) gebeurtenissen becommentariëren.

Methode

Voor deze studie werd een twittercorpus samengesteld met behulp van TWiNL (<https://twinl.surfsara.nl/>). Het corpus bevat alle tweets die gedurende 1 jaar (van 13/09/2015 tot 13/09/2016) gepost zijn met de hashtag 'dialect'. De motivatie voor het selecteren van '#dialect'-tweets ligt in het feit dat de hashtag reeds aangeeft dat de twitteraar actief reflecteert over het concept 'dialect'. Zoals hierboven aangegeven, functioneren die hashtags immers als een soort meta-commentaren op de tweet. Volgens de twitteraars zelf gaan de berichten in het corpus dus allemaal op de een of andere manier over het concept 'dialect'. In totaal leverde die dataverzameling een corpus van 735 tweets op. Na het verwijderen van dubbels en tweets die niet meer toegankelijk waren, bleef er een dataset van 493 analyseerbare tweets over. Elke tweet werd geannoteerd voor verschillende aspecten die samengevat worden in Tabel 1.

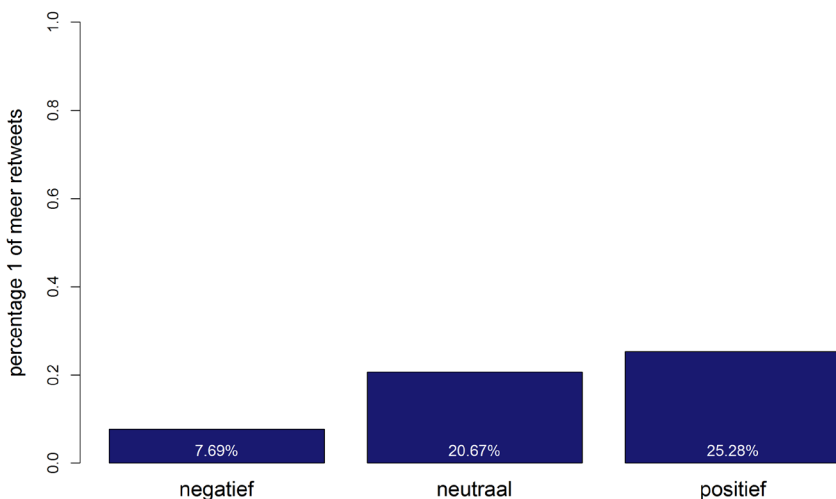
Annotatie	Betekenis
<i>Dialect</i> (1) Algemeen (2) Specifiek:Brugs, Gronings, West-Vlaams, ...	Wordt er over dialect in het algemeen getweet of over een specifiek dialect? In dat laatste geval, over welk dialect gaat het?
<i>Polariteit</i> (1) Positief (2) Negatief (3) Neutraal	Wordt dialect(gebruik) geëvalueerd? Zo ja, is het een positieve of negatieve evaluatie?
<i>Afkomst twitteraar</i> (1) België (2) Nederland	Komt de twitteraar uit België of Nederland (indien die informatie beschikbaar is op het openbare profiel van de twitteraar)?
<i>Gebruik dialect</i> (1) Ja (2) Nee	Wordt er dialect gebruikt in de tweet?
<i>Functie</i> (1) Informatieve tweets (2) Evaluatieve tweets (3) Expressie lokale identiteit (4) Dialect voor stilistisch effect (5) Observatie variatie	Wat is de functie van de tweet?

Tabel 1: Annotatie van het twittercorpus

Voor we verdergaan met een bespreking van de analyse van de tweets, lijkt het opportuun kort stil te staan bij de beperkingen van deze studie. Zoals aangegeven, is Twitter een interessante bron van informatie voor onderzoek naar cul-

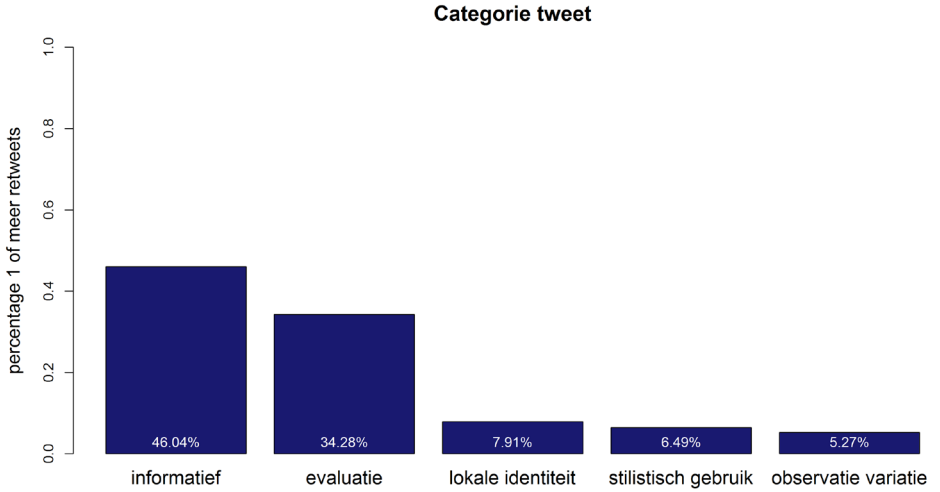
turele modellen, maar het medium brengt ook een aantal specificiteiten mee waarmee rekening gehouden moet worden bij de analyse. Ten eerste is er een strikte beperking op het aantal tekens in een tweet. Dat heeft ongetwijfeld invloed op wat en hoe er geschreven wordt op Twitter. Daarbij aansluitend willen we ook opmerken dat de korte lengte van twitterberichten en het ontbreken van enige vorm van context de interpretatie en, bijgevolg, analyse van de berichten in bepaalde gevallen bemoeilijkt. Tot slot zou ook het feit dat het om geschreven materiaal gaat een rol kunnen spelen: veel dialecten hebben geen geschreven vorm en zullen daarom misschien minder vaak terug te vinden zijn op een medium als Twitter.

Aantal retweets in functie van polariteit

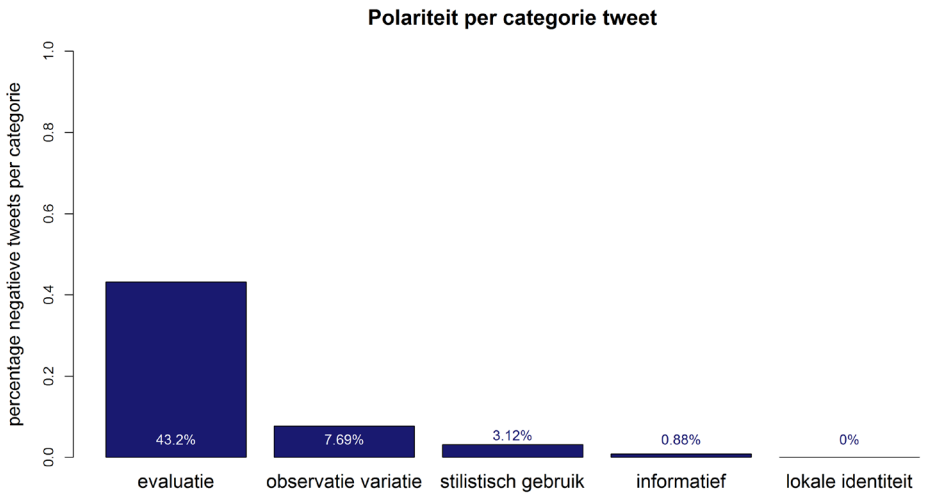


Figuur 1: Percentage tweets met minimaal 1 retweet in functie van polariteit

Over het algemeen blijkt dat de hashtag 'dialect' relatief weinig voorkomt in de tweets uit de onderzochte periode. Bovendien zijn er, van de 493 twitteraars die hem gebruiken, slechts 43 personen die hem meermaals tweeten. De twitteraars door wie de hashtag meer dan vijf keer gepost wordt, blijken, op basis van de informatie die ze verschaffen op hun twitterpagina, bijna allemaal (semi-) professioneel met taal bezig te zijn. De tweets lijken ook niet erg populair te zijn. De meeste '#dialect'-tweets krijgen geen of slechts één retweet en geen, één of twee likes, hoewel er natuurlijk ook tweets voorkomen die iets populairder zijn (het maximale aantal likes is 21, het maximale aantal retweets is 7). Opvallend



Figuur 2: Percentage tweets per functionele categorie



Figuur 3: Percentage negatieve tweets per functionele categorie

daarbij is wel dat er meer geretweet wordt als de boodschap die in de tweet wordt verwoord, een positieve of neutrale invalshoek heeft (zie Figuur 1).

Inhoud en polariteit van de '#dialect'-tweets

Om te bepalen welke culturele modellen kunnen afgeleid worden uit de '#dialect'-tweets, is het vooral interessant om de functionele categorisatie van de tweets te bekijken. Zoals aangegeven in Tabel 1, werden de tweets ingedeeld in vijf categorieën: (1) informatieve tweets, (2) tweets die te maken hebben met de evaluatie van een dialect, (3) tweets waarin de lokale identiteit van een dialectgebruiker in de verf gezet wordt, (4) tweets waarin een stilistisch effect beoogd wordt met het gebruik van of het tweeten over dialect en (5) tweets waarin de twitteraar de rijkdom aan variatie die voorkomt in de verschillende dialecten bespreekt. Voorbeelden (1) tot (5) geven voor elke categorie een voorbeeld uit het corpus.

(1) informatief

“Speciaal voor #Valentijnsdag: de liefste Limburgse koosnaampjes!
<http://www.eenofandereblog.nl/even-serieus/taal/limburglief-de-10-limburgse-liefkozingen/> ... #taal #dialect”

(2) evaluatie

“Ik zou gewoon al keihard in de lach schieten met Hanne’s dialect moest die mij sensueel in de oor spreken tijdens die proef #demol”

(3) lokale identiteit

“de brugseblogger zit vol ideeën !! drop hier u favoritiegugs woord !!!
#iedereenwestvlaams #brugge #dialect @StadBrugge”

(4) stilistisch gebruik

“Ne goeie vesse pistoleej #Arendonk #dialect #lekker”

(5) observatie variatie

“Collega’s hoe noemen jullie in je eigen #dialect: afrastering controleren? In Ommen zeggen we: vrèèn maakn of rikn”

Figuur 2 geeft de proportie tweets per categorie aan. Ongeveer 46% van de tweets behoren tot de eerste, informatieve categorie. Die categorie bevat voornamelijk reclametweets voor blogberichten over het dialect of voor evenementen in verband met een bepaald dialect. Dat sluit aan bij wat hierboven al werd opgemerkt: de meeste twitteraars die '#dialect' meer dan vijf keer gebruiken, zijn op een (semi-)professionele manier met taal bezig. Daarom is het niet verwonderlijk dat zij vooral informatieve berichten delen op Twitter. In iets meer dan 34% van de tweets wordt een dialect (of af en toe een dialectspreker) geëvalueerd of heeft de twitteraar het over de verstaanbaarheid of gepastheid van het dialect in een bepaalde context (categorie 2). De overige tweets behoren tot de categorieën 'lokale identiteit' (ongeveer 8%), 'stilistisch gebruik' (ongeveer 6%) of 'observatie variatie' (ongeveer 5%). Een aantal lokale-identiteitstweets bevatten ook een link naar een lied dat gezongen wordt in het dialect. Een verwijzing naar dat soort folklore is voor de twitteraar een manier bij uitstek om zijn lokale identiteit te benadrukken.

Verdere informatie over welke culturele modellen de twitteraars uitdragen, kan afgeleid worden uit de polariteit van de tweets. Tweets die een positieve boodschap verkondigen met de hashtag 'dialect', tonen immers aan dat de twitteraar positief staat tegenover het bestaan van het dialect en, bijgevolg, van taalvariatie (romantisch model). Negatieve tweets met de hashtag als metacommentaar wijzen er daarentegen op dat de twittergebruiker taalvariatie eerder storend vindt (rationalistisch model).

Over het algemeen blijken de meeste '#dialect'-tweets een positieve of neutrale invalshoek te hebben. Slechts 15.82% van de tweets zijn negatief polair. Zoals Figuur 3 aangeeft, komen negatieve tweets vooral voor bij de functionele categorie 'evaluatie'. Uit een verdere analyse van het corpus blijkt dat de tweets dan meestal een van de drie onderwerpen die in (6) tot (8) worden opgesomd, behandelen. De twitteraars uit het corpus gebruiken de hashtag dikwijls in een negatieve context als ze het hebben over communicatieproblemen die voortkomen uit het gebruik van een dialect. Daarnaast maken de negatieve tweets soms deel uit van een discussie waarin de status van een bepaald dialect of van een bepaalde taal (veelal het Fries) besproken wordt. Ten slotte valt het op dat de negatieve evaluatieve tweets meestal een specifiek dialect bespreken. Slechts in een handvol gevallen post men iets negatiefs over een spreker van het dialect (cf. voorbeelden 6-8).

(6) verstaanbaarheid of gepastheid van dialect in een bepaalde context

“Telefonisch interview met iemand die zwaar dialect spreekt (om ‘t tactisch te brengen), wat een uitdaging. Hoe vaak mag je ‘wablief’ herhalen?”

(7) status van een dialect of taal

“@BertRTV hahaha alleen dat dialect. ... nee ik zie fries niet als taal”

(8) evaluatie van een bepaald dialect

“2minuten Temptation gekeken en ik had al kotsneigingen van die Louise haar dialect.”

Conclusie

Over het algemeen kunnen we stellen dat uit het twittercorpus voornamelijk een romantisch model van taalvariatie naar voren komt. In slechts 34% van de tweets in het corpus wordt een rationalistische visie op taalvariatie teruggevonden. Dat zijn met name de tweets waarin de twitteraar een dialect of de verstaanbaarheid ervan evalueert. Voorbeeld (6), bijvoorbeeld, kan beschouwd worden als een tweet die duidelijk een rationalistische visie op taalvariatie uitdraagt.

Een romantische visie op het concept ‘dialect’ komt frequenter voor in het twittercorpus. Dat heeft misschien te maken met het feit dat er een bias zit in de groep gebruikers die de hashtag tweeten. Hierboven bleek bijvoorbeeld al dat de productievare twitteraars uit ons corpus, die de hashtag minstens vijf keer gebruiken, bijna allemaal op een (semi-)professionele manier met taal bezig zijn.

Ten eerste wordt de romantische visie uitgespeeld in de grootste groep tweets, namelijk de berichten met de functie om informatie over dialecten te verspreiden. Interesse voor de studie en het behoud van dialecten past immers in het romantische denkpatroon dat regionale taalvariatie als een rijkdom beschouwt. Verder kunnen ook de lokale-identiteitstweets en de tweets waarin over het dialect gesproken wordt om een stilistisch doel te verkrijgen, beschouwd worden als tweets waaruit een romantische visie op taalvariatie blijkt. Ook de tweets waarin een verwijzing naar een lied in het dialect gepost wordt, geven duidelijk een romantische visie op taalvariatie aan. Ten vierde wijst ook het feit dat slechts een minderheid van de tweets een negatieve invalshoek hebben erop dat mensen die ‘#dialect’ gebruiken, meestal een positieve houding hebben ten opzichte van

het dialect. Bijgevolg is hun visie op taalvariatie ook eerder romantisch dan rationalistisch: taalvariatie kan gebruikt worden als middel om je eigen identiteit te benadrukken.

Zoals aangehaald in de inleiding, kozen we voor de hashtag ‘dialect’ omdat we veronderstellen dat dat concept voor twitteraars bekend genoeg is en er dus ook relatief veel materiaal met die hashtag beschikbaar is. Hoewel een logische uitbreiding op deze studie zou zijn om ook andere hashtags in het onderzoek op te nemen, moeten we er rekening mee houden dat niet alle concepten bekend of saillant genoeg zijn voor taalkundige leken en dat ze daardoor relatief weinig voorkomen op Twitter. Een concept als ‘regiolect’, bijvoorbeeld, is wellicht niet gekend door twitteraars die geen taalkundige achtergrond hebben. Voor dergelijke concepten is het daardoor moeilijker om de *folk theories* over taal verder uit te werken. Toch hopen we in de toekomst ook andere concepten en hashtags te betrekken in dit onderzoek dan ‘#dialect’. Metacommentaren op andere concepten kunnen verdere interessante inzichten bieden in de *folk theories* van taal en ons helpen onderzoeken of we ook bewijs vinden voor dezelfde culturele modellen die een rol spelen in het denken over dialect.

Bibliografie

- Berthele, R. (2010). Investigations into the folk’s mental models of linguistic varieties. In D. Geeraerts, G. Kristiansen & Y. Peirsman (red.), *Advances in Cognitive Sociolinguistics*. Berlin, Boston: Mouton de Gruyter, 265-290.
- Dirven, R., H.-G. Wolf & F. Polzenhagen (2007). Cognitive Linguistics and Cultural Studies. In D. Geeraerts & H. Cuyckens (red.), *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford University Press, 1203-1221.
- Geeraerts, D. (2003). Cultural models of linguistic standardization. In R. Dirven, R. Frank & M. Pütz (red.), *Cognitive Models in Language and Thought. Ideology, Metaphors and Meanings*. Berlin: Mouton de Gruyter, 25-68.
- Holland, D. & N. Quinn (red.) (1987). *Cultural Models in Language and Thought*. Cambridge: Cambridge University Press.
- Polzenhagen, F. & R. Dirven (2008). Rationalist or romantic model in globalisation? In G. Kristiansen & R. Dirven (red.), *Cognitive Sociolinguistics. Language Variation, Cultural Models, Social Systems*. Berlin, Boston: Mouton de Gruyter, 237-300.

Isoglossen in Twitter? Op zoek naar tweet-geografie

Hans van Halteren en Roeland van Hout

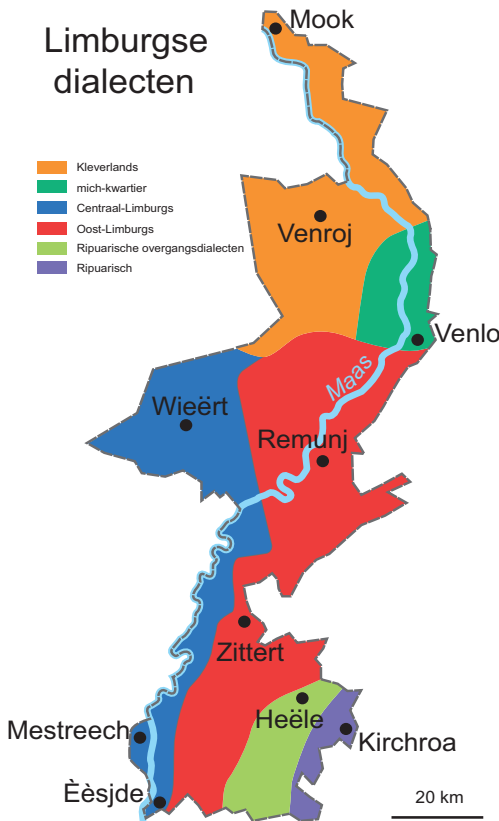
Inleiding¹

Op Twitter communiceren de gebruikers op allerlei manieren, waarbij ze zich in relatief korte boodschappen moeten manifesteren (max. 140 tekens). Tweets worden geschreven en het lijkt daarom voor de hand te liggen dat twitteraars in Nederland dat dan doen in de taal waarin ze hebben leren schrijven: het Nederlands. Misschien dat ze wat losse woorden in het dialect gebruiken, om hun tweets wat sjeu te geven, om hun boodschap wat kracht bij te zetten of misschien om te laten zien uit welke regio of plaats ze komen. Het gebruik van meer dialect lijkt bovendien pas aantrekkelijk als het nog intensief gebruikt wordt in de dagelijkse communicatie in de betreffende regio. Nu wordt in de provincie Limburg nog veel dialect gesproken. In de enquête van *Dagblad De Limburger* in 2016 (23 mei) geeft 72% van de respondenten aan nog dialect te spreken. Dat is een cijfer dat nergens anders in de Nederlandse regio's wordt gehaald. Wat Limburg verder interessant maakt, is het gegeven dat er binnen de provincie grote verschillen in dialect zijn, zoals bijvoorbeeld te zien is in Figuur 1 met een kaart van de Limburgse dialectgebieden. De Limburgse dialectgrenzen worden uitgebreid besproken in *Riek van Klank* (Keulen e.a. 2007).

Zo is in Figuur 1 de Uerdinger lijn te zien die onder Venlo doorloopt. Daarboven is het voornaamwoord van de eerste persoon enkelvoud *ik*, daaronder zeggen de dialectsprekers *ich*. Een andere isoglosse is de Panninger lijn, ten zuiden/oosten waarvan de initiële *s* voor een obstruent klinkt als *sj* (gepalatiseerd). De Uerdinger

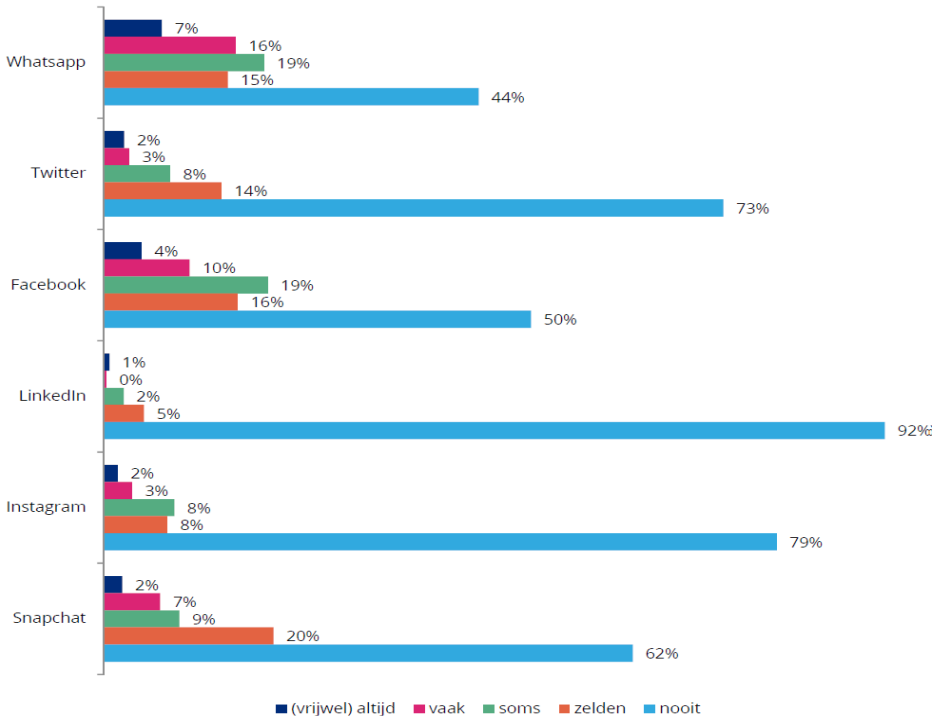
1 We danken Romy Roumans, die met haar werkstuk *Sjnoetebooch of Snoetebook? Herdefiniëren van de Panninger en Uerdinger lijn binnen de Limburgse dialectisoglossen aan de hand van Limburgse tweets* (voor de cursus Taalvariatie en de Sociale Media van de opleiding Taalwetenschap aan de RU Nijmegen) mede een aanzet heeft gegeven voor het hier beschreven onderzoek.

lijn scheidt het Kleverlands en het *mich*-kwartier van de zuidelijke Limburgse dialecten. De Panninger lijn doet dat deels ook, maar buigt bij Panningen naar het zuiden en scheidt het Centraal-Limburgs (rood) van het Oost-Limburgs (blauw). Het *mich*-kwartier (groen) omvat Venlo en een aantal plaatsen naar het noorden en westen toe en scheidt de zuidelijke objectsvorm *mich* van de noordelijke *mij(n)*-vormen. De *mich*-grens loopt dus noordelijker dan de *ich*-grens. Wat keert er van deze isoglossen terug in de tweets van de twitteraars? Kunnen klassieke dialectgrenzen teruggevonden worden in het taalgedrag in de moderne sociale media?



Figuur 1: De dialectgebieden in Nederlands Limburg; de plaatsnamen zijn in het Limburgs opgenomen (Cartografie: Amt für rheinische Landeskunde, Bonn, op verzoek van de Raod veur 't Limburgs)

In de bovengenoemde enquête is ook gevraagd naar het dialectgebruik op de sociale media. De resultaten staan in Figuur 2 die laat zien dat Limburgers op alle sociale media dialect gebruiken, maar in verschillende mate. Het gebruik van dialect is duidelijk het minst op LinkedIn, wat niet zo verwonderlijk is omdat dit medium beroepsgericht is. De media waar meer dialect gebruikt wordt zijn Whatsapp en Facebook, die blijkbaar wat directer, persoonlijker en informeler worden gevonden dan Twitter. De verschillen in gebruik van dialect in de verschillende media weerspiegelen, zo lijkt het, het verschil in status en de gebruiksmogelijkheden van deze communicatieplatforms.



Figuur 2: De mate van gebruik van Limburgs in verschillende sociale media

Als 37% van de Limburgers die twitteren wel eens of vaker dialect gebruikt, dan moeten daar de sporen van zijn terug te vinden in het enorme bestand aan tweets dat ondertussen beschikbaar is (het TwiNL corpus, verzameld door het Dutch eScience Centre; zie Tjong Kim Sang & Van den Bosch 2013). In deze bijdrage willen we die sporen terugvinden en vaststellen in hoeverre het dialectgebruik dat we op Twitter aantreffen overeenkomt met wat we weten over de Limburgse dialecten. We richten we ons op de klassieke isoglossen. Vinden we die terug in het tweetgedrag van de Limburgers? Het gaat om de Uerdinger lijn, de Pan-ninge lijn, het *mich*-kwartier (met de bijbehorende *mich*-lijn) en om de *dich*-lijn met *dich*-vormen in het zuiden en *ou/oe*-vormen in het noorden. Deze isoglosse verloopt deels iets noordelijker dan de *mich*-lijn (zie Bakker 2016).

Methodie

De eerste stap was om het dialect terug te vinden in de berg van tweets die beschikbaar is voor onderzoek. Hiervoor hebben we eerst de twittergebruikers afkomstig uit Nederlands Limburg geïdentificeerd. Dat kunnen we uiteraard niet doen aan de hand van hun dialectgebruik (ook al om te voorkomen dat we vinden wat we willen vinden), want Limburgers wonen ook buiten Limburg en niet alle Limburgers tweeten in het dialect. Gelukkig biedt Twitter zelf een oplossing. Naast de tekst van de tweets die je normaal op het scherm ziet, stuurt Twitter ook metadata mee, zoals het tijdstip waarop ze verstuurd zijn, en mits beschikbaar (d.w.z. als de verzender toegang gegeven heeft tot zijn GPS), ook de geografische positie vanwaar de tweet is verzonden. Door te zoeken naar gebruikers die een groot deel (wij kozen voor 2/3) van hun tweets vanuit eenzelfde gemeente verstuurd, hoopten we de gebruikers te koppelen aan hun woonplaats. Dit systeem is niet waterdicht. Door het criterium van het grootste deel van de tweets uit dezelfde gemeente (2/3) zorgen we ervoor dat het grootste deel van de twitteraars thuishoren in de gemeente waar we ze plaatsen. Verder is door de mobiliteit in het moderne Nederland de kans groot dat mensen niet meer in de plaats wonen waar ze zijn opgegroeid en waar ze dus (wellicht) hun dialect hebben geleerd. We kunnen dus niet claimen dat de twitteraar is opgegroeid in de plaats van waaruit hij/zij twittert. We vonden 13.221 sprekers die a) aan een gemeente konden worden toegewezen, b) tussen begin 2011 en half 2015 minstens 100 tweets / 1.000 woorden aan originele tweets (dus geen retweets) hadden geproduceerd, en c) hoofdzakelijk in het Nederlands schreven (dit met name om interferentie van Duitse twitteraars te ondervangen).

Resultaten

Een zorg vooraf was of er wel voldoende dialect gevonden zou worden in onze twittergroep. De samplingsprocedure bij de aanleg van het twittercorpus is potentieel zelfs tegen het selecteren van stevige dialectgebruikers gericht. Zo gaat de verzameling van TwiNL uit van het vinden van tweets aan de hand van Standaardnederlandse woorden, en eist onze geografische toewijzing dat de twitteraar beschikt over een GPS-vaardig invoermechanisme (vaak dus smartphone). Toch bleek deze zorg snel overbodig. We zochten naar de vormen *ich*, *mich* en *dich*, inclusief spellingsvarianten met *ie* i.p.v. *i* en *g* i.p.v. *ch*, alsmede naar vormen met palatalisatie (vooraan het woord een *s* of *z*, gevolgd door *j*, dan een *l*, *m*, *n*, *p*, *t* en daarna nog een klinker *a*, *e*, *i*, *o*, *u*). Dit leverde de aantallen gebruikers (twitteraars) en de gebruiksfrequentie (aantal voorkomens) op van Tabel 1.

Dialectvorm	Aantal gebruikers	Percentage Gebruikers	Aantal voorkomens
<i>ich</i>	6.069	45,9	113.705
<i>mich</i>	4.370	33,1	40.005
<i>dich</i>	5.254	39,7	64.904
<i>sj-</i>	2.468	18,7	13.404

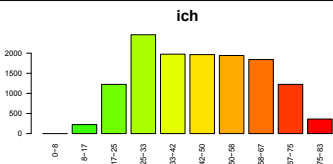
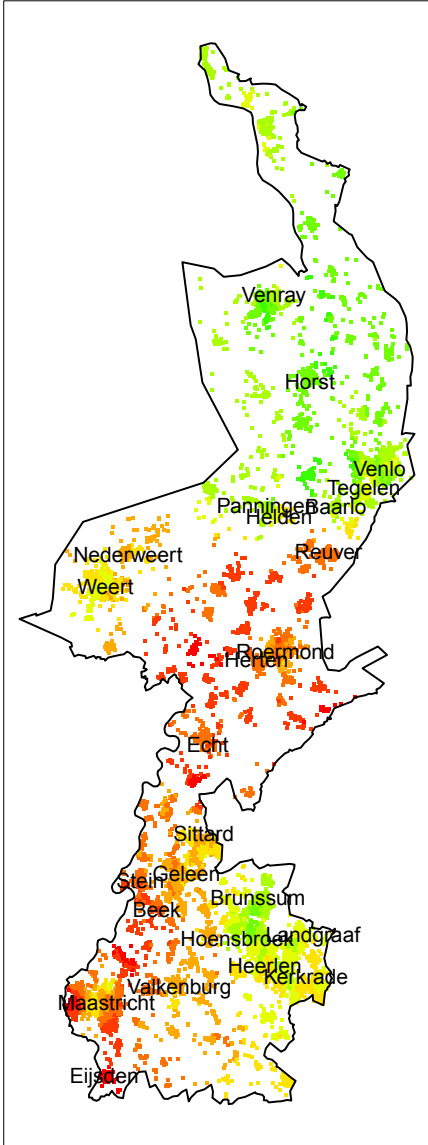
Tabel 1: Limburgse gebruikers en gebruik van vier dialectverschijnselen

Een aanzienlijk percentage van de twitteraars gebruikte de vormen, en aan de aantallen voorkomens is te zien dat dit ook niet slechts een enkele keer gebeurde. Bovendien zagen we aan de voorbeelden dat er ook daadwerkelijk sprake was van dialectgebruik. Veel tweets zijn geheel in dialect:

- (1) *Iech weet neet wie iech hub gelegen, mer gweet wel dat iech neet good hub gelegen! Auw miene ruk!* [Maastricht]
- (2) *ig weit nog sjteeds neet wat t versjil is. daarom, doe snaps toch waal detse neet kens doorbuize asse vnv zous gaon :P* [Reuver]
- (3) *det zuut d'r good oet! Kald beej he! De vastelaovend waas veur mich snel afgelaupe heb weer kaelontstaeking* [Venlo]

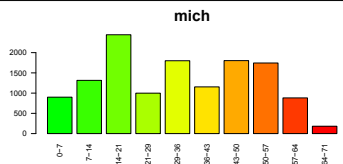
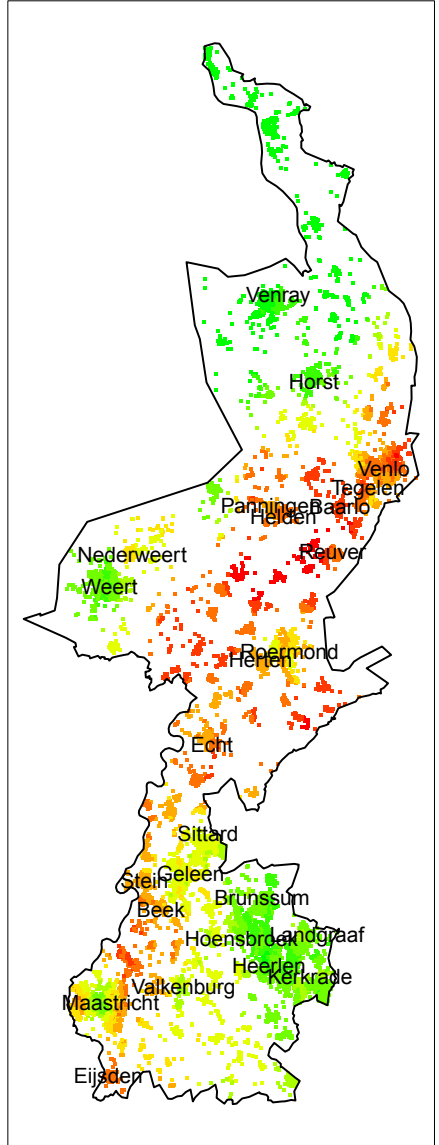
We zien in bovenstaande voorbeelden ook hoe zeer de spelling kan variëren, inclusief het aan elkaar schrijven van voornaamwoorden en het werkwoord (*gweet*). Helaas kunnen we die aaneengeschreven vormen voorlopig niet goed uit de tweets lichten. Twitteraars mengen ook het dialect, niet alleen met het

ich



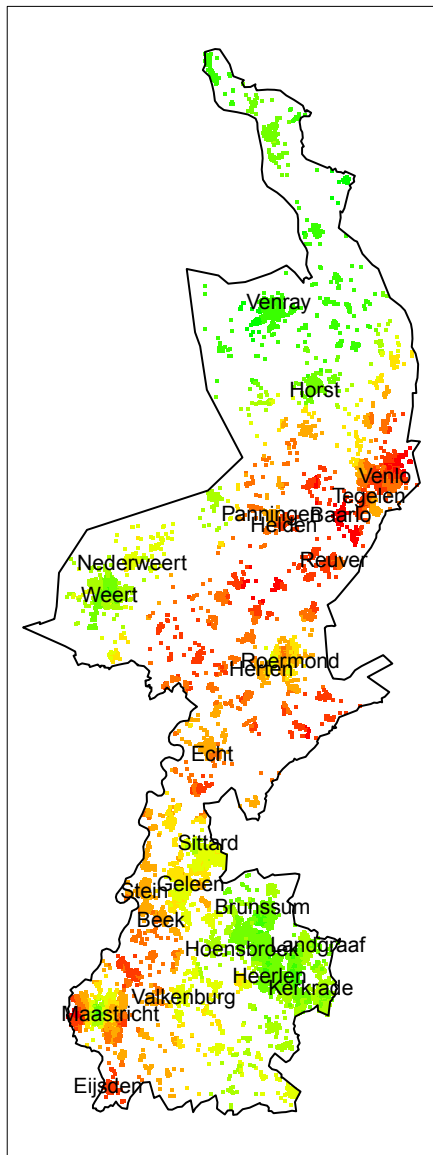
Figuur 3: Het gebruik van ich

mich

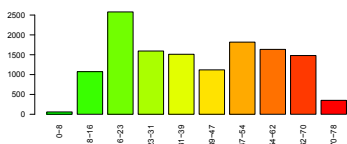


Figuur 4: Het gebruik van mich

dich

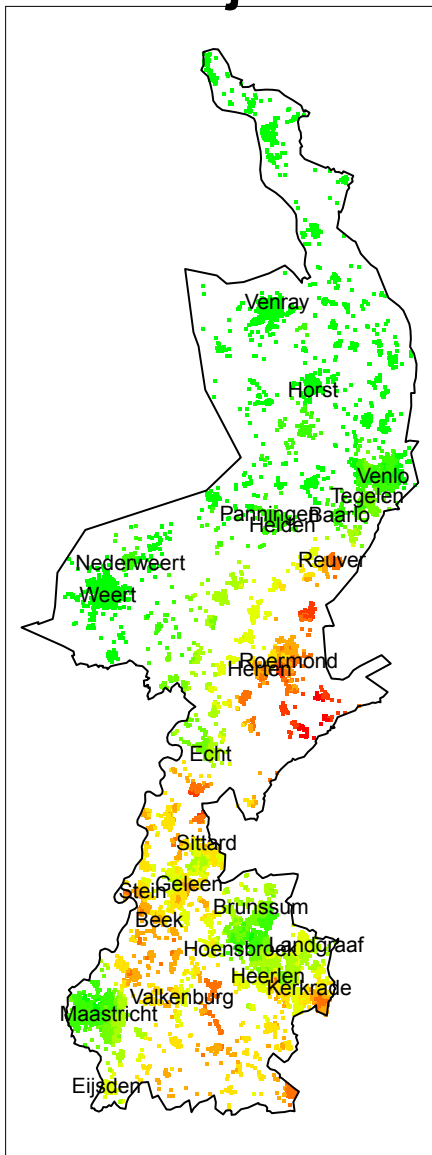


dich

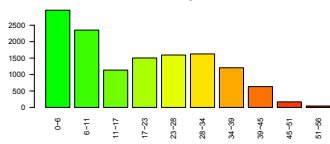


Figuur 5: Het gebruik van dich

sj



sj



Figuur 6: Het gebruik van sj-klank

Nederlands, maar uiteraard ook, modern als ze zijn, met het Engels:

(4) *volgesmig is mien account gewist ofzo, hea zeat det ich spotify free gebroek terwiel ich unlimited hub.*

(5) *download dich dit spelke: coin dozer echt vet verslavend.*

Gezien de gevonden aantallen konden we overgaan tot het in kaart brengen van de geografische verspreiding van het gebruik van de gezochte vormen. De vier kaarten zijn te vinden in Figuren 3–6. Elke kaart toont alle 13.221 sprekers in ons onderzoek en geeft de plaatsnamen van alle plaatsen waar minstens 100 sprekers in onze steekproef zitten. De kleur van elk punt geeft het percentage vormgebruik aan dat rond die spreker kan worden waargenomen, berekend op basis van de 100 dichtstbijzijnde sprekers (inclusief de spreker in kwestie zelf). De kleur verloopt van groen voor 0% naar rood voor het hoogste waargenomen percentage voor de betreffende vorm. Dat laatste varieert per verschijnsel. Bij elke kaart staat een legenda die de kleuren koppelt aan de percentages bij die vorm. We hebben de legenda als histogram afgebeeld, om meteen een overzicht te krijgen van het totale aantal gebruikers. Hoe hoger hoe meer gebruikers uit die subgroep. Het bij rood horende percentage wijzigt met de vorm, zodat de kaarten in dat opzicht niet rechtstreeks vergelijkbaar zijn; de kleuren moeten eerder gelezen worden als maat voor relatief gebruik.

De *ich*-kaart bevat fraaie resultaten. De groene kleuren in het noorden van Limburg worden geel bij Venlo, zoals de Uerdinger lijn voorspelt. Weert en Nederweert laten het gebruik van de *ich*-vorm zien, maar Nederweert wat sterker, wellicht omdat daar meer dialect wordt gebruikt. Maastricht kleurt rood, met geel in het midden, hetgeen aangeeft dat Maastricht een interne geografische differentiatie kent, met minder dialectgebruik in het centrum. Ook in Venlo zien we geografische differentiatie, wat erop wijst dat dat dialectgrenzen in Twitter een wat diffuser karakter hebben, mede ten gevolge van de mobiliteit. Verder zijn Venlo (*ik*) en Tegelen (*ich*) in de loop van de tijd aaneen gegroeid. De Oostelijke Mijnstreek is duidelijker minder rood gekleurd, het vertoont zelfs groene plekken, hetgeen we niet verwacht hadden. Het dialectgebruik lijkt daar minder intensief dan in andere delen van Limburg. Globaal genomen kunnen we in elk geval vaststellen dat de Uerdinger lijn terugkeert in de Limburgse twitteraars en hun tweets.

Hoe is het met de *mich*-kaart? Daar zou in elk geval het *mich*-kwartier zichtbaar moeten worden. Venlo kleurt rood en de gele vlekken naar het noorden

en oosten geven netjes het *mich*-kwartier aan. Horst valt duidelijk buiten het *mich*-gebied. Maastricht laat weer interne differentiatie zien, met zelfs wat groen, wijzend op een laag aandeel van dialect in de twitteraars in het centrum. Opvallend is de groene kleur van de Oostelijke Mijnstreek, wederom wijzend op een laag gebruiks-aandeel van het dialect. Ook Weert is groen gekleurd, maar Nederweert niet.

De kaarten van *mich* en *dich* lijken sterk op elkaar. Het belangrijkste verschil is vlak onder Horst te vinden. Horst kleurt groen en dat is ook de verwachting omdat het buiten het *dich*-gebied valt. De plaatsen onder Horst zijn op de *dich*-kaart wat meer oranje gekleurd dan op de *mich*-kaart en dat klopt met het verloop van de *doe+dich*-isoglosse (zie voor de details Bakker 2017). Kortom de *dich*-kaart illustreert dat onze aanpak gevoelig genoeg is om subtiele dialect-geografische verschillen aan het daglicht te brengen.

De *sj*-kaart laat zien dat in Venlo en noordelijk van Venlo het aandeel *sj*-gebruikers zeer laag is. Onder Panningen worden de kleuren geler, maar niet in het gebied van Weert en Nederweert. Dat komt fraai overeen met de loop van de Panninger lijn. De plaats Echt is groen, evenals Maastricht, wat wederom het verloop van de Panninger lijn weerspiegelt. De smalle strook die langs de Maas in een groene kleur zou moeten lopen (zie Figuur 1) zien we niet terug. Een reden kan zijn dat de plaatsen te weinig twitteraars hebben, in combinatie met het gegeven dat lang niet alle Limburgers in de plaats wonen waar ze geboren zijn. Migraties en verhuizingen dragen bij aan het mengen van groepen dialectsprekers, ook aan de vestiging van niet-Limburgers in Limburg. Rond Roermond worden de meeste twitteraars met de palatale *sj*- aangetroffen. En wederom is het gebruik van de *sj*- in de Oostelijke Mijnstreek laag, met name in Hoensbroek en Brunssum.

Conclusies

De geografie van Twitter heeft al eerder interessante resultaten laten zien op woordniveau. Een bekend voorbeeld is de keuze tussen *patat* (noordelijk) en *frites* (zuiden) (Stroop 2005). De scheidslijn is de Maas (<http://www.hpdetijd.nl/2015-08-26/potato-potahto-het-taalkundige-onderscheid-tussen-patat-en-friet-geduid/>). Dit verschil in de keuze tussen *frites* en *patat* laat zien hoe nieuwere vormen van taalvariatie zich weerspiegelen in de ruimte, zoals nieuwe vormen van taalvariatie ook weer een tijdsdimensie of een sociale inbedding hebben. Die complexe inbedding geldt natuurlijk ook voor de oudere vormen

van taalvariatie, de vormen met een geschiedenis die zich weerspiegelt in de dialecten.

In deze bijdrage hebben we ons gericht op het gebruik van dialectvormen in Limburg, vormen met een verleden. We hebben daarbij de ingang gekozen van dialectwoorden en hun varianten, waarbij we de twitteraars als uitgangspunt kozen in de verzameling van de data, gebruikmakend van de informatie over hun geografische positie. De dataverzameling is zeker niet perfect, daar zijn we van overtuigd, maar de grote hoeveelheden compenseren de gebreken. We gaan het materiaal dat we vermeld hebben verder doorzoeken om meerdere dialectkenmerken in kaart te brengen. We willen ook in meer detail de geografische reikwijdte van woorden afbakenen en die informatie weer gebruiken voor een diepere analyse van dialectgebruik.

De conclusie waarmee we willen eindigen, is vooral dat de klassieke dialectgrenzen in de tweetgeografie terugkeren, vaak wat diffuser, maar dat is begrijpelijk gezien de mobiliteit van dialectsprekers, de invloed van de standaardtaal op het dialect en het gegeven dat veel mensen in het Nederlands twitteren, ook in Limburg. Het lijkt erop dat we juist ook plaatsen in het vizier krijgen waar het dialectgebruik minder is. Dat geldt voor Weert en dat geldt voor de oostelijke mijnstreek. Een opvallend resultaat tot slot was de interne geografische differentiatie die onze methode oplevert in grotere plaatsen zoals Maastricht en Venlo. Behalve dat Twitter een aantrekkelijke dataleverancier kan zijn in brede zin heeft het ook nog het voordeel dat we gegevens hebben die relatief spontaan zijn, niet gesproken weliswaar maar geschreven. De communicatie op Twitter is duidelijk minder normgebonden dan het traditionele schrijfverkeer.

Bibliografie

- Bakker, F. (2017). *Waar scheiden de dialecten in Noord-Limburg? Een dialectometrisch onderzoek naar het gewicht van isoglossen*. Proefschrift Radboud Universiteit Nijmegen. LOT Dissertation 435.
- Keulen, R., Wijngaard, T. van de, Crompvoets, H. & Walraven, F. (2007). *Riek van klank, Inleiding in de Limburgse dialecten*. Sittard: Veldeke (Veldeke Taalstudies nr. 2).
- Stroop, J. (2005). Wat u moet weten voor u gaat eten. In V. De Tier, R. Keulen & J. Swanenberg (red.) *Proeven van dialect. Een kijkje in de keuken van de Nederlandse dialecten*. Groesbeek: Stichting Nederlandse Dialecten, 15-35.

Tjong Kim Sang, E. & van den Bosch, A. (2013). Dealing with big data: The case of Twitter. *Computational Linguistics in the Netherlands Journal* 3, 121-134.

Hun twitteren: Tweets als bron voor onderzoek naar syntactische taalvariatie

Stefan Grondelaers, Roeland van Hout en Hans van Halteren

Een basisprobleem in onderzoek naar (morfo)syntactische variatie, is het feit dat alternerende constructies in een tekst of gesprek veel minder voorkomen dan bijvoorbeeld woorden en klanken. Daardoor is er niet zelden een gebrek aan voorbeelden (tokens) in de databestanden (de corpora) die voor taalkundig onderzoek gebruikt kunnen worden (Milroy & Gordon 2003). Vaak zijn er te weinig gegevens om de precieze rol in kaart te brengen van de factoren die het variatieproces sturen. Dat frequentieprobleem wordt nog nijpender wanneer het onderwerp van onderzoek een gestigmatiseerde non-standaardvariant is die gebruikers bewust afgeleerd wordt, zoals het gebruik van *hun* als onderwerpsvorm in het Nederlands.

De typische bronnen die taalkundigen als corpusmateriaal hanteren, bieden in het geval van dat gestigmatiseerde *hun* geen uitkomst: of je nu 10 miljoen of 10 miljard woorden krantentaal verzamelt, geen enkele redacteur zal het *hun* (dat hij in zijn gesproken taal misschien gebruikt) ook schrijven, laat staan dat er een eindredacteur is die het zou laten staan. Om het als subject gebruikte *hun* in een corpusanalyse te onderzoeken is er taal materiaal nodig waarin het normbesef van gewone geschreven taal een beperktere rol speelt. Dat is het geval voor internetchat, maar een minstens even interessante bron die in elk geval het frequentieprobleem oplost is het gigantische twittercorpus TwiNL dat het Meertens Instituut verzameld heeft (Tjong Kim Sang & Van den Bosch 2013). Tweets bevatten ook veel non-standaardvormen, maar een mogelijk nadeel is dat hun formaat (maximum 140 tekens) misschien te kort is voor onderzoek naar syntactische variatie. Vanuit sociolinguïstisch perspectief komt daar meteen de vraag bij over het soort Nederlands dat in tweets te vinden is: het is uiteraard geschreven taal, maar de inhoud en stijl is vaak erg informeel en er zijn duidelijke kenmerken van spreektaal in te vinden.

Al eerder is vastgesteld (cf. van Hout 2006) dat *hun* zich razendsnel verspreidt als onderwerpsvorm (zie 2), op posities waar de onderwerpsvorm juist *zij* had moeten zijn (zie 1):

- (1) Als je zo speelt, krijgen zij natuurlijk altijd kansen. (standaard)
- (2) Als je zo speelt, krijgen hun natuurlijk altijd kansen.

Overigens speelt dit veranderingsproces zich uitsluitend in het Nederlandse Nederlands af. Het is categorisch afwezig in het Belgische Nederlands. De oudste bekende observatie van het fenomeen is van 1911 (Vor der Hake 1911), maar er is sprake van een hoorbare toename in de laatste decennia. Het Corpus Gesproken Nederlands bevat 213 voorkomens van *hun* tegenover 1048 voorkomens van *zij*. Deze voorkomens van *hun* worden door een klein aantal sprekers (<20 %) in spontane, onvoorbereide face-to-face-gesprekken of telefoonconversaties geproduceerd.

Hun rukt onstuitbaar op, zo lijkt het, en daar zijn een aantal redenen voor aan te wijzen. Die hebben we al eerder onderzocht, maar daar hebben we telkens andere technieken voor nodig, die niet gemakkelijk in één geïntegreerde analyse samen te brengen zijn (Grondelaers, Van Gent & Van Hout te verschijnen). Tot nu hebben we gevonden dat de onderwerpsvorm *hun* zowel linguïstische als sociale betekenis heeft:

- > Linguïstisch: in een reeks corpusanalyses hebben we gevonden dat *hun* duidelijk geprefereerd wordt boven *zij* in contexten van ‘geëngageerd negatief contrast’ (“Wij zijn Ajax, hun moeten oprotten”);
- > Sociaal: in een reeks sprekerevaluatie-experimenten vonden we dat *hun* nieuw, dynamisch prestige uitdrukt. Door *hun* te gebruiken profileren sprekers zich als uitdagend, assertief, energetisch, cool ...

Het feit dat de onderwerpsvorm *hun* die nieuw-prestigieuze sociale betekenis heeft, wijst erop dat in het Nederlandse Nederlands belangrijke ideologische veranderingen aan de gang zijn: de opkomst van *hun* belichaamt de ondergang van de uitgesproken conservatieve standaardtaalideologie die één variëteit – uniform Standaardnederlands – als enige correcte, zuivere, mooie variëteit ziet. Het feit dat *hun* prestige kan krijgen – zelfs al gaat het om nieuw prestige – suggereert het bestaan van een parallel waardensysteem met een eigen hiërarchie die vormen legitimeert die de conservatieve ideologie uitsluit.

Twitter zou niet alleen een interessante databron kunnen zijn om het frequentieprobleem te omzeilen, maar ook om de linguïstische en sociale betekenis van *hun* in één geïntegreerde analyse te onderzoeken. Twitter is beschikbaar in enorme hoeveelheden (> 2 miljoen/dag – Tjong Kim Sang 2011). Tweets kunnen gekarakteriseerd worden als informele geschreven taal met kenmerken van gesproken taal, waardoor ze zeer geschikt zijn als “supplementary data for investigating non-frequent, non-canonical phenomena in spoken language” (Rehbein 2014: 20).

We hebben tweets met de onderwerpsvorm *hun* of de onderwerpsvorm *zij* (in het meervoud) uit het TwiNL corpus geselecteerd (Tjong Kim Sang & Van den Bosch 2013). Om er zeker van te zijn dat we goede voorbeelden hadden, waarin *hun* als subject gebruikt werd, en niet als object of bezittelijk voornaamwoord, hebben we onze extractie beperkt tot niet-ambigue patronen waarin *hun* of *zij* aan het begin van de tweet of zin staan, en voor frequente persoonsvormen (zoals *hun zijn* en *zij zijn*). Het resultaat was 14.658 tweets, met een evenwichtige verdeling tussen *hun* (7.514) en *zij* (7.144), waarin de nieuwe vorm dus een licht overwicht heeft. Dat is een hoeveelheid waarmee we aan de slag konden om onderzoek te doen naar factoren die de keuze tussen beide voornaamwoorden bepalen.

Er is nog maar weinig onderzoek naar de linguïstische kenmerken van Twitter gedaan, maar een aantal dingen valt onmiddellijk op. Er zijn veel (erg) korte boodschappen (gewoonlijk korter dan het maximum van 140 tekens), veel vaste uitdrukkingen, geritualiseerde frases (*zij/hun zijn goed* komt maar liefst 776 keer voor, plus nog 209 keer met spelvarianten van *goed*), identieke tweets en massa's retweets. Verder bevatten 2.404 tweets alternatieve spellingen, spellingen waarvan het bedoeling is dat ze afwijken:

- (3) *hun zijn gewooooooon geweldiggg*
- (4) *hun zijn go**ho**edd*

Wat is het voordeel van alternatief spellen? Je jaagt er in elk geval de taalpurist bijna even hoog mee op de kast als met de onderwerpsvorm *hun*, maar ze geven ook iets aan. Tweets (3) en (4) zijn duidelijke voorbeelden van orthografische intonatie. Er kan ook sprake zijn van een economisch motief wanneer het om verkorte vormen gaat:

- (5) zij zijn **gwn** stoer
- (6) m'n broer had **t** toch tegen m'n ouders gezegd

Maar je kunt je tweet ook alternatief spellen om je zelf als een dynamisch, non-conformistisch, schalks persoon neer te zetten:

- (7) chinezen kijken niet zo nauw meer: zij denken dat de schulden **p**roblematiek van tijdelijke **aald** is

De eigen regels en de lagere normgebondenheid van Twitter lijken in het voordeel te werken van het gebruik van de onderwerpsvorm *hun*.

Om uit te zoeken welke factoren een rol spelen in de keuze tussen *hun* en *zij*, zijn alle tweets bekeken en gecodeerd voor een reeks van kenmerken die we gekozen hebben in functie van twee hypothesen. In vroeger werk van Helen de Hoop (Van Bergen e. a. 2011) is gesuggereerd dat *hun* rijker van betekenis is dan *zij* omdat het uitsluitend naar menselijke referenten verwijst (terwijl *zij* ook naar stoffelijke referenten kan verwijzen); we noemen dat vermoeden de animacy-hypothese. Om na te gaan welk effect deze factor op de voorkeur voor *zij* of *hun* heeft hebben we de referenten van *hun* en *zij* allemaal voor de factor menselijkheid gecodeerd. In de tweets verwijzen beide vormen overigens zelden naar stoffelijke referenten, 31 keer (0.2%). Beide onderwerpsvormen zijn dus echte persoonlijke voornaamwoorden. In die 31 voorkomens wordt 1 keer *hun* gebruikt en 30 keer *zij*, maar dat frequentieverschil kan niet het verschil in de gebruiksfrequenties verklaren tussen *hun* en *zij* in de overige 14.727 voorkomens. Het is zelfs opvallend te noemen dat *hun* blijkbaar toch naar stoffelijke referenten kan verwijzen. In ons eigen werk (Grondelaers, Van Gent & Van Hout te verschijnen) hebben we aangetoond dat *hun* vooral wordt ingezet om contrast, en bij voorkeur dan nog negatief contrast uit te drukken. Dat wordt geïllustreerd in de volgende voorbeelden:

- (8) suriname mensen zijn best wel dom en hun denken dat wij in nederland dom zijn. #fail en denk logisch na man
- (9) wat hun denken is fictie. wat wij hebben is een #feit
- (10) huh! hun doen lootjes trekken bij kerstmis, en wij bij sinterklaas! fett! #glee

We hebben dat eerder de contrasthypothese genoemd. Als *hun* een geëngageerde contrasteerder is, dan verschijnt het wellicht vaker in tweets waarin een contrast

wordt aangezet en versterkt met intensiveerders, tussenwerpsels of niet-talige symbolen. Voorbeelden van intensiveerders zijn:

- (11) hun zijn tering veatt!!
 (12) <@> hahahha schatje niks tegen homo's hoor maar hun doen kankersneu #loveyou

Een voorbeeld van een tussenwerpsel is:

- (13) kijk wat hun doen wtf #rtl5

De niet-talige symbolen zijn vooral allerlei visuele symbolen die allerlei emoties kunnen verbeelden.

De voorspelling klopt voor de drie contrastmarkeerders. De resultaten voor deze drie contrastmarkeerders staan in Tabel 1.

		<i>hun</i>	<i>zij</i>
Intensiveerders	geen	4028 40,0%	6039 60,0%
	één of meer	3486 75,9%	1110 24,1%
Tussenwerpsels	geen	5784 46,8%	6568 53,2%
	één of meer	1730 75,5%	576 24,5%
Niet-talige symbolen	geen	6606 49,8%	6655 50,2%
	één of meer	903 64,9%	488 35,1%

Tabel 1: De verdeling van hun en zij over drie soorten van contrastmarkeringen; de optelsom van de voorkomens kan minder zijn dan 14.658 door ontbrekende coderingen (omdat het onmogelijk bleek een code toe te kennen)

De vorm *zij* in Tabel 1 is ver in de minderheid in de tweets met een of meer intensiveerders. We vinden dat effect in vergelijkbare mate voor tussenwerpsels en in een duidelijk wat zwakkere mate voor niet-talige symbolen, maar het effect is er. Wanneer een verschil, een contrast wordt aangezet of benadrukt, komt de vorm *hun* vaker voor.

Het is natuurlijk de vraag of taalgebruikers zich bewust zijn van de contrastbetekenis van *hun*. Dat kunnen we verder onderzoeken door naar de spelling van de tweets te kijken. We zeiden al dat 2404 tweets een alternatieve spelling bevatten. Voor de spelling in tweets hebben we een driedeling gemaakt in “geen afwijking”, “alternatieve spelling” (opzettelijk afwijkend) en “spelfouten” (onbedoeld incorrect gespeld). De kwalificatie alternatieve spelling hebben we alleen toegekend aan die gevallen waar de twitteraar met duidelijke opzet afwijkt van de standaardspelling. Twijfelgevallen zijn onder spelfout geschoven. Het resultaat staat in Tabel 2.

Spelling	<i>hun</i>	<i>zij</i>
Geen afwijking	5748 47.2%	6432 52.8%
Alternatieve spelling	1264 77.4%	369 22.6%
Spelfout	459 59.5%	312 40.5%

Tabel 2: De verdeling van *hun* en *zij* over drie spellingscategorieën

Hun is duidelijk in de meerderheid in tweets met een alternatieve spelling. *Hun* wordt dus bevorderd door een foute, maar vooral door een bewust foute spelling.

Hoe werken contrastmarkeringen en spelling samen? Hier vonden we een zeer opvallend patroon. De contrastmarkeringen doen feilloos hun werk in de correct gespelde tweets, in de zin dat ze het gebruik van *hun* in onze statistische modellen (veel) voorspelbaarder maken. Aan tweets met alternatieve spelling voegen ze opvallend genoeg helemaal niets toe: in die tweets spelen contrastmarkeringen geen enkele rol. De intensiveerders en tussenwerpsels zijn overigens zelf vaak al bewust fout gespeld:

- > Intensiveerders: *zooo*, *errug*, ...
- > Tussenwerpsels: *focking*, *facking*, *vukking*, *fett*, *kloowte*, ...

Dat opvallende onderscheid kunnen we alleen verklaren als we ervan uitgaan dat er in de alternatief gespelde tweets bewust gestileerd wordt, dat de twitteraar zichzelf via taal wil profileren als een coole, dynamische, provocerende Nederlander. De sociale betekenis van *hun* is blijkbaar voldoende bekend om bewust ingezet te kunnen worden in zo'n dynamische identiteitsprofilering.

De vorm *hun* wordt met andere woorden op twee niveaus gebruikt: als onbewust gehanteerde pronomenvariant, en als bewust gebruikt stileringsmiddel. In beide gevallen ligt de dynamische sociale betekenis aan de basis van het gebruik. We zouden nog verder kunnen gaan door te suggereren dat de linguïstische betekenis van *hun* (negatieve contrastering) en de sociale betekenis (de *hun*-gebruiker is een dynamische, provocerende Nederlander) deels overlappen: uitdaging en provocatie veronderstellen een derde partij die het object is van dat antagonistische gedrag. Het animacy-effect dat Van Bergen e.a. (2011) vinden – het feit dat *hun* vaker dan *zij* naar mensen verwijst – kan in hetzelfde licht verklaard worden. Het omgekeerde kan echter niet: als *hun* als enige voordeel zijn exclusief menselijke verwijzing heeft, kan je niet verklaren waarom al de contrasteffecten die we vinden optreden.

De uitgangsvraag van deze bijdrage kunnen we in elk geval positief beantwoorden: Twitter is een buitengewoon rijke databron voor onderzoek naar (morfo-) syntactische variatie. Er is gigantisch veel materiaal, er blijken voldoende tokens te vinden van niet- (of nog niet)-standaardvarianten, en de maximale lengte van 140 karakters is duidelijk geen belemmering om de factoren te identificeren die variatie aandrijven. Twitter bevat bovendien voldoende ‘onderbuik’ om sociale aandrijvers in beeld te brengen naast de linguïstische. Om al die redenen wacht Twitter ongetwijfeld een grote toekomst als databron voor heel veel soorten taalwetenschappers.

Bibliografie

- Bergen, G. van, W. Stoop, J. Vogels & H. de Hoop (2011). Leve Hun! Waarom hun nog steeds hun zeggen. *Nederlandse Taalkunde* 16, 2–29.
- Grondelaers, S., P. van Gent & R. van Hout (te verschijnen). On the inevitability of social meaning and ideology in accounts of syntactic change. Evidence from pronoun competition in Netherlandic Dutch. In T. Christensen & T. Jensen (red.), *Explanations in Sociosyntax. Dialogue across paradigms*. Amsterdam/Philadelphia: Benjamins.
- Hout, R. van (2006). Onstuitbaar en onuitstaanbaar. Ontstaan en verspreiding van het onderwerp “hun”. In N. van der Sijs, J. Stroop, F. Weerman (red.), *Wat iedereen van het Nederlands met weten en waarom*. Amsterdam: Bert Bakker.
- Milroy, L. & Gordon, M. (2003). *Sociolinguistics: Method and Interpretation*. Oxford: Blackwell.

- Tjong Kim Sang, E. (2011). Het gebruik van Twitter voor taalkundig onderzoek. *Tabu* 39, 62-72.
- Tjong Kim Sang, E. & van den Bosch, A. (2013). Dealing with big data: The case of Twitter. *Computational Linguistics in the Netherlands Journal* 3, 121-134.
- Vor der Hake, J. (1911). Is de beleefdheidsvorm U'n verbastering van UEd.? *De Nieuwe Taalgids* 5, 16-34.

