

Ronde 4

Roland de Bonth (a) & Cefas van Rossem (b)
 (a) Instituut voor de Nederlandse Taal, Leiden
 (b) Olympus College, Arnhem; Meertens Instituut, Amsterdam
 Contact: roland.debonth@ivdnt.org
 c.vanrossem@olympuscollege.nl

Krant in de klas 2.0

1. Inleiding

Kranten vormen een bijzonder interessante bron voor taalkundig, letterkundig en historisch onderzoek. Al vanaf de zeventiende eeuw worden er in de Lage Landen kranten gedrukt, waarin de lezers kennis konden nemen van belangrijke nationale en internationale politieke en militaire gebeurtenissen, zoals ‘staatsbezoeken’, ‘veldslagen’ en ‘vredesonderhandelingen’. Daarnaast vormden kranten een ideaal medium om pas verschenen of te verschijnen boeken aan te kondigen, maar ook om verloren ringen, weggelopen honden en gevluchte dieven onder de aandacht van de lezers te brengen.

2. Delpher

De Koninklijke Bibliotheek in Den Haag heeft de afgelopen jaren ruim twee miljoen kranten van 1618 tot en met 1995 gescand, en deze foto's via de website 'Delpher.nl' gratis beschikbaar gesteld. Daaronder bevinden zich niet alleen dagbladen die in Nederland zijn gedrukt, maar ook Nederlandstalige kranten uit Nederlands-Indië, de Antillen, de Verenigde Staten en Suriname.

De scans van de kranten uit Delpher zijn, met behulp van *Optical Character Recognition*, omgezet in voor de computer leesbare tekst. Zo is het mogelijk geworden om daarin te zoeken naar specifieke woorden, combinaties van woorden en (historische) spellingvarianten.

In de kranten zijn heel verschillende teksten te bekijken. Denk hierbij maar eens aan de verschillende manieren waarop geadverteerd kon worden. Advertenties voor diensten

en producten komen voor in allerlei perioden en in vele vormen. Nieuwsberichten omvatten niet alleen de verslagen van het dagelijkse nieuws, zoals we dat nu nog steeds kennen, maar bijvoorbeeld ook de aankondigingen van vertrekkende of arriverende schepen. Behalve taalkundig onderzoek – bijvoorbeeld naar woorden die eindigen op het suffix *-ette* – is het dus ook zeer goed mogelijk om historisch of sociaal onderzoek te doen rond speciale gebeurtenissen of veranderende gewoonten.

3. Couranten Corpus

Een verhaal apart vormen de zeventiende-eeuwse kranten. De kwaliteit van de automatisch herkende tekst bleek door de kwaliteit van het toenmalige drukwerk (en daarmee van de afbeeldingen), het gebruikte (gotische) lettertype en de afwijkende taal van onvoldoende kwaliteit om er onderzoek mee te doen. Daarom hebben ruim 200 vrijwilligers, onder leiding van Nicoline van der Sijs, duizenden exemplaren van 13 verschillende zeventiende-eeuwse kranten die op Delpher te vinden zijn, zorgvuldig getranscribeerd via een op het Meertens Instituut ontwikkelde *crowdsourcing*-omgeving. In 2020 ontving deze dataset de Nederlandse dataprijs.

De metadata van deze handmatig overgetikte artikelen zijn door stagiairs en medewerkers van het Instituut voor de Nederlandse Taal (INT) gecorrigeerd, waarbij alle artikelen eigen metadata hebben gekregen en dubbele kranten uit de dataset zijn verwijderd. Ook zijn nieuwe metadata toegevoegd, bijvoorbeeld over ‘genre’ en ‘nieuwssoort’ (advertenties, binnenlands nieuws, buitenlands nieuws, enz.). De bijna 19 miljoen woorden zijn automatisch taalkundig verrijkt met modern lemma en woordsoort, zodat ze beter vindbaar zouden zijn. Het historisch lexicon van het INT – waarin historische spellingvarianten zijn opgenomen – maakt het mogelijk dat het aantal zoekresultaten zonder al te veel moeite sterk toeneemt.

De tot nu toe getranscribeerde zeventiende-eeuwse kranten bevatten 109.532 artikelen en zij zijn alle opgenomen in het zogeheten ‘Couranten Corpus’. De informatie in deze kranten is interessant voor onderzoekers van verschillende disciplines: van historici en historisch taalkundigen tot literatuurwetenschappers en kunsthistorici. Overigens wordt op dit moment – wederom onder leiding van Nicoline van der Sijs – met behulp van ‘Transkribus’ gewerkt aan de ontsluiting van nog meer zeventiende-eeuwse kranten.

4. Onderwijs en onderzoek

De kranten uit Delpher, waaronder die exemplaren die zijn opgenomen in het Couranten Corpus, bevatten een schat aan gegevens voor (wetenschappelijk) onderzoek, maar zij kunnen ook heel goed worden ingezet bij onderwijs aan leerlingen en studenten.

Belangrijk is daarbij wel dat je goed moet weten wat je precies wilt onderzoeken (*Wat is je hoofdvraag?*), hoe je dat kunt onderzoeken (*Wat zijn je zoektermen?*) en hoe je moet zoeken (*Wat zijn je zoekopdrachten?*).

Cefas van Rossem, docent Nederlands aan het Olympus College in Arnhem en gast-onderzoeker van het Meertens Instituut, zal aan de hand van enkele voorbeelden laten zien hoe leerlingen en studenten de mogelijkheden van Delpher kunnen benutten voor een (taalkundig) onderzoek met betrekking tot het ‘*Virgin Islands Creole Dutch*’.

Roland de Bonth, onderzoeker bij het Instituut voor de Nederlandse Taal en daarvoor twintig jaar docent Nederlands in het voortgezet onderwijs, zal, aan de hand van het Couranten Corpus, met behulp van enkele voorbeelden, uitleggen wat voor onderzoek je met dit corpus kunt uitvoeren, hoe je optimaal gebruik kunt maken van de zoekmogelijkheden van het corpus en op welke wijze de gevonden resultaten verwerkt kunnen worden.

5. Opbrengst

Na het volgen van deze presentatie zal duidelijk zijn geworden dat ‘Delpher’ en ‘het Couranten Corpus’ twee bronnen zijn die zich heel goed lenen voor onderzoek in de klas naar sterk uiteenlopende onderwerpen, en dat slim zoeken zorgt voor een optimaal resultaat.

Referenties

- Bonth, R. de & C. van Rossem (2021-nu). “Achter het Achtervoegel’, maandelijkse bijdrage over het suffix -ette”. In: *Neerlandistiek.nl*. Online raadpleegbaar op: www.neerlandistiek.nl.
- Couranten Corpus (2023). Online raadpleegbaar op: <https://couranten.ivdnt.org/corpus-frontend/couranten/search/>.
- Delpher (2023). Online raadpleegbaar op: <https://www.delpher.nl>.
- Sanders, E. (2011). *Eerste Hulp Bij e-Onderzoek, voor studenten in de geesteswetenschappen, Slimmer zoeken, slimmer documenteren*. Amsterdam/Den Haag/Leiden: Early Dutch Books Online.
- Sanders, E. (2015). *Digitaal goud delven bij de KB (Koninklijke Bibliotheek)*. Den Haag: Koninklijke Bibliotheek.