

Submission 43

**Het *Getuigenissen*-project als digitalisering van historisch taalkundig onderzoek**

Magda Serwaczak

*Vrije Universiteit Brussel, Brussel, België*

In het *Getuigenissen*-project worden getuigenverklaringen en verhoren van verdachten uit achttiende- en negentiende-eeuwse criminele rechtbanken in Vlaanderen onder de loep genomen (<https://www.getuigenissen.org>). De documenten worden door een team van vrijwilligers en deskundigen getranscribeerd en vervolgens in een elektronisch corpus opgenomen. Als resultaat kunnen ze op grote schaal onderzocht worden door o.a. taalkundigen, sociologen, criminologen en historici. Voor historisch sociolinguïsten zijn deze bronnen bijzonder waardevol, want ze bieden een – al zij het onrechtstreekse – weerspiegeling van de gesproken omgangstaal van vrijwel alle sociale klassen. Gezien de duidelijke verschillen tussen spreek- en schrijftaal ook vandaag de dag, kunnen we deze bronnen echter niet beschouwen als een rechtstreekse weergave van gesproken taalgebruik op schrift (cf. Chafe & Tannen, 1987; Kytö & Walker, 2003). Dit leidt tot de vraag in hoeverre de geschreven bronnen een betrouwbare weergave van zijn de gesproken interactie en welke sporen van de spreektaal erin gevonden kunnen worden? Deze bijdrage belicht dus enkele aspecten van de relatie tussen spreektaal en schrijftaal in de documenten uit het *Getuigenissen*-corpus. Aan de hand van een kwantitatieve data-analyse in R wordt de frequentie van geselecteerde markeerders van oraliteit en schriftelijkheid berekend (Biber, 1988; Rutten & van der Wal, 2014) en worden onderliggende gebruikspatronen opgespoord. Vervolgens wordt een multivariate regressieanalyse uitgevoerd om mogelijke interacties tussen onafhankelijke variabelen (waaronder leeftijd en gender) te ontrafelen. Als laatste stap worden kwalitatieve onderzoeksmethodes toegepast om meer licht te werpen op het taalgebruik van individuele sprekers. Deze bijdrage streeft een dubbele doelstelling na. Enerzijds willen we de mogelijkheden belichten die de digitalisering van historische bronnen en de toepassing van methodes uit de digitale geestwetenschappen binnen de historische sociolinguïstiek met zich meebrengen. Anderzijds trachten we om een gedetailleerd beeld te schetsen van de manieren waarop gesproken taal in schriftelijke bronnen wordt gereconstrueerd.