



Universiteit
Leiden
The Netherlands

Verfransing onder de loep: Nederlands-Frans taalcontact (1500-1900) vanuit historisch-sociolinguïstisch perspectief

Assendelft, B.M.E.

Citation

Assendelft, B. M. E. (2023, May 24). *Verfransing onder de loep: Nederlands-Frans taalcontact (1500-1900) vanuit historisch-sociolinguïstisch perspectief*. LOT dissertation series. LOT, Amsterdam. Retrieved from <https://hdl.handle.net/1887/3618471>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3618471>

Note: To cite this publication please use the final published version (if applicable).

Hoofdstuk 4 Corpus en methode

4.1 Inleiding

Dit hoofdstuk behandelt de totstandkoming van het *Language of Leiden*-corpus oftewel *LOL*-corpus, dat is samengesteld om de invloed van het Frans op het Nederlands te kunnen onderzoeken. De naam van het corpus verwijst naar het feit dat het onderzoek zich specifiek richt op de taal van de Hollandse stad Leiden. Er is voor gekozen om het onderzoek tot Leiden te beperken om een aantal redenen. Ten eerste kan op deze manier invloed van regionale variatie worden beperkt. Rutten, Vosters & Van der Wal (2015) hebben de variabele regio betrokken in hun onderzoek naar het gebruik van Franse leensuffixen in achttiende- en negentiende-eeuwse brieven en vonden dat het gebruik van Franse suffixen lager werd naarmate de regio verder van de Nederlands-Franse taalgrens lag. De stad Amsterdam (die als aparte regio was meegenomen in het onderzoek¹⁴) vormde een uitzondering op deze trend: ondanks zijn ligging, relatief ver van het Franse taalgebied, was er een hoge frequentie in het gebruik van Franse suffixen te zien (Rutten, Vosters & Van der Wal 2015: 159–160). Taalcontactsituaties verschillen dus per locatie, en aangezien de focus in het huidige onderzoek ligt op de sociolinguïstische dimensie en niet op de regionale dimensie, is het beter om slechts één locatie in het onderzoek te betrekken.

De keuze voor Leiden als deze ene locatie heeft vooral met de geschiedenis van de stad te maken, en met name zijn textielindustrie en de migratiestromen die deze veroorzaakte. Zoals in het vorige hoofdstuk naar voren kwam, trokken met name van eind zestiende tot eind zeventiende eeuw veel immigranten uit de Zuidelijke Nederlanden naar Leiden, waaronder ook Franstaligen of tweetaligen. Leiden is dus de stad bij uitstek waar veel Franse invloed moet zijn geweest.

Een methodologische reden om het onderzoek te richten op Leiden heeft met de beschikbaarheid van materiaal te maken. Het

¹⁴ In het *Brieven als Buit*-corpus, dat voor dit onderzoek is gebruikt, is de stad Amsterdam om geografische en demografische redenen als aparte regio opgenomen en niet als deel van Noord-Holland. Geografisch gezien was Amsterdam gescheiden van de rest van Noord-Holland door water; demografisch gezien was Amsterdam in vergelijking met het omliggende land sterk geurbaniseerd, en trok de stad veel immigranten aan. Te verwachten was dat dit het Amsterdamse taalgebruik beïnvloedde (Rutten & Van der Wal 2014: 12).

stadsarchief van Leiden (Erfgoed Leiden en Omstreken) bevat een zeer uitgebreide en diverse collectie documenten uit de eeuwen die deel zijn van dit onderzoek. De beschikbare archieven representeren veel verschillende domeinen van de historische Leidse samenleving, zoals de kerken, de gilden en het privéleven. De aanwezigheid van documenten uit diverse sociale domeinen was cruciaal aangezien ‘domein’ een van de twee variabelen is die het uitgangspunt vormen van de opbouw van het *LOL*-corpus (zie verder paragraaf 4.2). De ruime hoeveelheid beschikbaar materiaal gaf aanleiding te vermoeden dat er voldoende teksten zouden zijn om een geschikt corpus mee op te bouwen.

Naast domein is er nog een tweede onafhankelijke variabele die de basis vormt van het corpus: periode. Er zijn acht verschillende perioden onderscheiden in het corpus: 1500–1549, 1550–1599, 1600–1649, 1650–1699, 1700–1749, 1750–1799, 1800–1849 en 1850–1899, zodat de diachrone ontwikkelingen van de invloed van het Frans op het Nederlands onderzocht kunnen worden. De zeven domeinen die zijn opgenomen in het *LOL*-corpus – Academie, Economie, Liefdadigheid, Literatuur, Privé, Publieke Opinie en Religie – trachten samen een representatieve afspiegeling te geven van de historische situatie van Leiden. Per periode en per domein is er gestreefd naar een woorden-aantal van 5.000, wat betekent dat het corpus idealiter 280.000 woorden bevat (8 perioden x 7 domeinen x 5.000 woorden). Voor enkele combinaties van periode en domein was het echter niet mogelijk om geschikt corpusmateriaal te vinden, waardoor het uiteindelijke corpus ongeveer 250.000 woorden bedraagt (zie paragrafen 4.3.1 en 4.3.3).

In de volgende paragrafen wordt uiteengezet hoe het *LOL*-corpus is opgebouwd en welke procedures zijn gevolgd bij de corpusopbouw. Paragraaf 4.2 gaat dieper in op de variabelen periode (4.2.1) en domein (4.2.2). Het *LOL*-corpus is het resultaat van archiefonderzoek, waarin relevante documenten zijn opgespoord, gefotografeerd en vervolgens getranscribeerd, om de teksten geschikt te maken voor de taalkundige analyses. Paragraaf 4.3 behandelt de verschillende stappen die zijn doorlopen tijdens de corpusopbouw en geeft per domein aan hoe de zoek- en selectieprocedure van de bronnen in zijn werk ging en welke problemen ik tegenkwam tijdens het archiefonderzoek (4.3.1). 4.3.2 behandelt de procedures en conventies die zijn gevolgd bij het transcriberen van alle documenten. Het hoofdstuk sluit af met een overzicht van de grootte en de indeling van het *LOL*-corpus als geheel (4.3.3) en een korte uiteenzetting van de

analyses die met dit corpus zijn uitgevoerd en die in de volgende hoofdstukken worden besproken (4.4).

4.2 De variabelen

Het *LOL*-corpus is ontworpen op basis van twee onafhankelijke variabelen: periode en domein. In paragraaf 4.2.1 wordt een toelichting gegeven op de variabele periode en paragraaf 4.2.2 gaat nader in op de zeven domeinen die het corpus vormen.

4.2.1 Periode

Het totale corpus beslaat vierhonderd jaar, beginnend bij het jaar 1500 en eindigend bij het jaar 1899. Zoals besproken in hoofdstuk 2 gaat het taalcontact tussen het Frans en het Nederlands echter veel verder terug dan het jaar 1500: er waren gedurende de middeleeuwen al handelscontacten en het Frans was vanaf de veertiende eeuw ook een administratieve taal in de Nederlanden. Het corpus begint dus niet bij de aanvang van het Nederlands-Franse taalcontact. De keuze om het corpus bij het jaar 1500 te laten beginnen heeft daarentegen te maken met de aard van dit onderzoek: er is amper geschikt materiaal beschikbaar van voor 1500 voor het soort corpusonderzoek dat in de huidige studie wordt uitgevoerd. Het bronnenmateriaal van voor de zestiende eeuw is veelal beperkt tot administratieve teksten en representeert daarmee slechts een deel van de domeinen die in het huidige onderzoek zijn meegenomen. Overigens is dit gebrek aan materiaal ook nog zichtbaar in de eerste helft van de zestiende eeuw: zoals in paragraaf 4.3.1 wordt beschreven, is voor vijf van de zeven domeinen die het *LOL*-corpus omvat geen materiaal uit de eerste helft van de zestiende eeuw gevonden.

Een andere reden om het onderzoek in de zestiende eeuw te laten beginnen, is het feit dat in de loop van de zestiende eeuw het gebruik van het Frans in de Nederlanden intensiverde door de komst van Franstalige religieuze vluchtelingen uit de Zuidelijke Nederlanden (zie hoofdstuk 2 en 3). Tegelijkertijd ving in de zestiende eeuw de kritiek op de invloed van het Frans op het Nederlands aan. Dit blijkt uit de publicatie van puristische woordenboeken in deze tijd. Aangezien het belang van het Frans als *lingua franca* in Europa in de loop van de negentiende eeuw langzaam maar zeker afnam en plaatsmaakte voor het

Duits en het Engels als belangrijke internationale talen, loopt de onderzochte periode in dit proefschrift tot en met 1899.

Om de diachrone ontwikkelingen van veranderingen in het Nederlands onder invloed van het Frans te kunnen achterhalen, is de totale periode van vierhonderd jaar die het corpus omvat opgedeeld in acht perioden van vijftig jaar: 1500–1549, 1550–1599, 1600–1649, 1650–1699, 1700–1749, 1750–1799, 1800–1849 en 1850–1899. Op deze manier kan de invloed van het Frans op het Nederlands in deze acht perioden met elkaar vergeleken worden. Er is gekozen voor perioden van vijftig jaar omdat het gebruik van kortere perioden (bijvoorbeeld tien of twintig jaar) te weinig materiaal per tijdvak zou opleveren. Ook andere historisch-sociolinguïstische corpora die taalveranderingen diachroon onderzoeken zijn opgedeeld in tijdvakken van vijftig jaar; dit geldt voor het *ARCHER*-corpus en voor *GerManC* (zie hoofdstuk 2). Nevalainen & Raumolin-Brunberg (2016: 29–30) geven in hun bespreking van het *CEEC*-corpus (zie hoofdstuk 2) aan dat zij meestal hebben gekozen voor perioden van veertig jaar (in plaats van hun oorspronkelijke verdeling in perioden van twintig jaar), omdat perioden van veertig jaar betrouwbaardere resultaten gaven met betrekking tot correlaties en ontwikkelingen dan kortere perioden. Ook hieruit blijkt dus dat grotere tijdvakken de voorkeur verdienen vanwege de beperkte beschikbaarheid van materiaal.

Voor elk van de acht onderscheiden perioden is gestreefd naar een totaal aantal woorden van 35.000, gelijkmatig verdeeld over de zeven domeinen (zie paragraaf 4.2.2 hierna).

Om een goede diachrone vergelijking te kunnen maken, is het van belang dat de documenten die voor elk van de perioden worden geselecteerd, een goede representatie zijn van de periode en dat de dateringen ervan ook substantieel afwijken van die in de periode ervoor en erna. Zo is het ongeschikt om bijvoorbeeld voor de periode 1750–1799 documenten te nemen die alleen in het laatste decennium van deze periode zijn geschreven en om tegelijkertijd voor de periode 1800–1849 alleen teksten uit het eerste decennium te nemen, omdat de documenten in de tijd dan erg dicht bij elkaar liggen, terwijl het taalgebruik in de teksten toch verschillende perioden moet representeren. Om dit te voorkomen, is een extra criterium opgesteld: alle documenten die een bepaalde periode representeren, moesten ofwel uit het midden van de periode afkomstig zijn (dus voor de periode 1750–1799 rond 1775), ofwel moesten de documenten verspreid over de gehele periode geschreven zijn. Dit betekent dat er niet alleen documenten uit de jaren

1810 gebruikt konden worden, maar dat er dan bijvoorbeeld ook documenten uit de jaren 1830 in het corpus moesten worden opgenomen. Dit criterium moest ervoor zorgen dat er genoeg tijd tussen de documenten uit verschillende perioden zou zitten.

4.2.2 Sociaal domein

In hoofdstuk 2 is al ingegaan op het concept ‘domein’ en het verschil tussen een domein en een genre. Domeinonderscheidingen zijn bij uitstek geschikt om taalcontactsituaties te onderzoeken; vandaar dat er voor het *LOL*-corpus, waarmee een specifieke taalcontactsituatie wordt onderzocht, te weten het taalcontact tussen het Nederlands en het Frans in de historische stad Leiden, is gekozen voor domeinonderscheidingen. Door domeinen te onderscheiden kon bovendien een corpus worden gebouwd dat representatief is voor de geschiedenis van Leiden. Er zijn voor het corpus zeven domeinen geselecteerd: Academie, Economie, Liefdadigheid, Literatuur, Privéleven, Publieke Opinie en Religie. De domeinen zijn gekozen op basis van algemene domeinen die worden onderscheiden in de literatuur, zoals economie/werk, privéleven/familie en vrienden, publieke opinie/media, religie/kerk (zie de hoofddomeinen die Fishman 1968 en Greenfield 1968 onderscheiden en de studie van Kessels-Van der Heijde 2002 naar taalkeuze in Maastricht, die in hoofdstuk 2 zijn besproken), aangevuld met enkele domeinen die specifiek in Leiden een belangrijke rol speelden (Academie, Liefdadigheid, Literatuur). In hoofdstuk 3 is het belang van elk van de gekozen domeinen in de geschiedenis van Leiden ook al toegelicht. De keuze voor deze domeinen komt grotendeels overeen met onderscheidingen die in studies over de geschiedenis van Leiden worden gemaakt. In deel 2 van *Leiden. De geschiedenis van een Hollandse stad* (Van Maanen & Groenveld 2003) zijn er bijvoorbeeld aparte hoofdstukken voor ‘het economische leven’, ‘sociale verhoudingen en maatschappelijke zorg’, ‘kerk en samenleving’, ‘het sociaal-culturele leven’ en ‘onderwijs en wetenschap’. Daarnaast bespreekt Tjalsma (1978) de structuur van Leiden in het midden van de achttiende eeuw in termen van de volgende ‘sferen’: woonsfeer, werksfeer, verzorgende sfeer, sfeer van de sociale zorg, overheidssfeer, kerkelijke sfeer, universitaire sfeer. Deze indelingen komen dus redelijk overeen met de domeinen die ik aanhoud. Een domein dat wel vaak wordt onderscheiden maar dat ik niet als apart domein heb opgenomen in mijn onderzoek, is het openbaar bestuur.

Zoals bij de beschrijvingen van de verschillende domeinen in hoofdstuk 3 al naar voren is gekomen, was er gedurende lange tijd sprake van verregaande bemoeienis van het stadsbestuur met onder andere het bedrijfsleven, het bestuur van de universiteit en met de kerkenraden: in de diverse besturen zaten vaak afgevaardigden van het stadsbestuur en beslissingen moesten vaak door de burgemeesters of het Gerecht van Leiden worden genomen. Het openbaar bestuur speelt dus een rol in verschillende domeinen en is daarmee indirect vertegenwoordigd in het corpus.

4.3 Procedures corpusopbouw

Voor elk van de zeven domeinen die zijn geselecteerd voor het *Language of Leiden*-corpus, is materiaal verzameld door middel van archiefonderzoek, zowel digitaal (in online databases en inventarissen) als fysiek (in de archieven zelf). Aangezien het corpus zich richt op de taal van Leiden, is de collectie van het archief van Leiden (Erfgoed Leiden en Omstreken) als uitgangspunt genomen; voor de domeinen Economie, Liefdadigheid en Religie was het archief Erfgoed Leiden en Omstreken zelfs de enige bron van het materiaal dat uiteindelijk in het corpus is opgenomen. Voor het domein Privé zijn naast het stadsarchief van Leiden ook de collecties van de Universiteitsbibliotheek van Leiden geraadpleegd, alsmede die van het Nationaal Archief te Den Haag en het Stadsarchief Amsterdam. Het materiaal voor het domein Literatuur komt ook deels uit de Leidse Universiteitsbibliotheek en uit de collecties van enkele andere archieven. Het grootste deel van het materiaal dat voor het domein Publieke Opinie is verzameld, komt van Delpher.nl. Het Academisch domein is het enige domein waar geen van de teksten uit het archief Erfgoed Leiden en Omstreken komt, maar waar de speciale collecties van de Universiteitsbibliotheek van Leiden de bron van al het corpusmateriaal zijn geweest.

Hieronder ga ik in op de dataverzameling, waarbij eerst enkele algemene principes aan de orde komen, en daarna per domein wordt uitgelegd hoe de dataverzameling in zijn werk is gegaan (4.3.1). De transcriptieconventies en de transcriptieprocedure komen aan bod in 4.3.2, en in 4.3.3 wordt een volledig overzicht van het *LOL*-corpus gegeven.

4.3.1 *Dataverzameling*

Bij het zoeken naar materiaal zijn enkele principes gevolgd. Zoals in paragraaf 4.2.1 al werd genoemd, was het noodzakelijk dat de documenten die voor een bepaalde periode zijn gebruikt, uit het midden van de betreffende periode komen of gelijkmatig over de periode verspreid zijn om zo te voorkómen dat de documenten uit twee opeenvolgende periodes dicht bij elkaar liggen.

Daarnaast was het van belang om ervoor te zorgen dat de individuele voorkeuren van één schrijver niet te veel invloed zouden hebben in het corpus. Daarom werd er een limiet aangehouden van 1.250 woorden per schrijver per domein per periode. Dit betekent dat er minimaal vier verschillende documenten, geschreven door verschillende auteurs, per periode per domein in het corpus moesten worden opgenomen om het vastgestelde woordenaantal van 5.000 per domein per periode te bereiken.

Ten derde moest de auteur een (sterke) band hebben met Leiden en er een aanzienlijke tijd hebben gewoond, aangezien het corpus de taal van Leiden representeert. Dit criterium was met name van belang bij het Privédomein en het domein Literatuur: bij de documenten die voor deze domeinen gebruikt zijn, kon vastgesteld worden wie de auteur van het betreffende document was, terwijl dat voor de overige domeinen meestal niet mogelijk was. Zoals bij de toelichting op beide domeinen hieronder zal blijken, is per individuele auteur bepaald hoe dit criterium is geïnterpreteerd en wat de overwegingen waren om een bepaald persoon als Leidenaar te rekenen.

Tot slot is per domein een keuze gemaakt voor een of (in het geval van het domein Economie) twee genres die de domeinen in het corpus representeren. Deze beperking is toegepast om ervoor te zorgen dat het materiaal dat een specifiek domein moet representeren zo uniform mogelijk is binnen één periode en over de verschillende perioden heen (zie ook Van de Voorde 2022: 66). De keuze voor een bepaald genre was gebaseerd op de beschikbaarheid van materiaal: er moesten genoeg documenten van het gekozen genre aanwezig zijn gedurende een zo groot mogelijke periode tussen 1500 en 1900 (zie verder de uitleg bij de verschillende domeinen hieronder).

Hieronder wordt per domein aangegeven hoe het archiefonderzoek precies in zijn werk is gegaan, welke problemen er ontstonden en welke keuzes zijn gemaakt. Ook is voor een indruk van het materiaal

voor elk domein een foto van een van de gebruikte teksten geselecteerd en samen met de bijbehorende transcriptie toegevoegd.¹⁵

Academie

Nederlandse teksten voor het Academische domein zijn schaars, omdat veel teksten in het Latijn zijn geschreven. Er zijn wel wetenschappelijke artikelen te vinden die in het Nederlands gesteld zijn, maar het is de vraag of deze als Leidse teksten gezien kunnen worden, aangezien de hoogleraren die deze schreven meestal maar kort aangesteld waren in Leiden en vaak ook niet oorspronkelijk uit de buurt van Leiden kwamen. Ik heb er daarom voor gekozen notulen van vergaderingen van het college van curatoren (en burgemeesters) te gebruiken. Ten eerste is hier voor elk jaar vanaf de oprichting van de universiteit materiaal van beschikbaar, en ten tweede waren de secretaris en hoogstwaarschijnlijk ook de klerken, die de notulen maakten, uit Leiden afkomstig (Roberts & Otterspeer 2003: 203–204), waardoor deze notulen als Leids kunnen worden aangemerkt.

De notulen worden alle bewaard bij de afdeling Bijzondere Collecties in de Universiteitsbibliotheek Leiden. Voor de gehele periode zijn de notulen handgeschreven en in banden van één of soms enkele jaren samengevoegd. Voor het overgrote deel van de handen is niet bekend van wie ze zijn, maar het verschil ertussen is doorgaans wel gemakkelijk te maken. Om aan het principe ‘niet meer dan 1.250 woorden per schrijver’ te voldoen, zijn daarom voor elke periode notulen in ten minste vier verschillende handen uitgezocht; daarnaast is ervoor gezorgd dat deze over de gehele periode verspreid zijn. Voor een deel van de eerste twee decennia van de zeventiende eeuw meldde de inventarisbeschrijving van de Universiteitsbibliotheek dat hier alleen afschriften van beschikbaar zijn; de betreffende inventarisnummers zijn daarom overgeslagen.

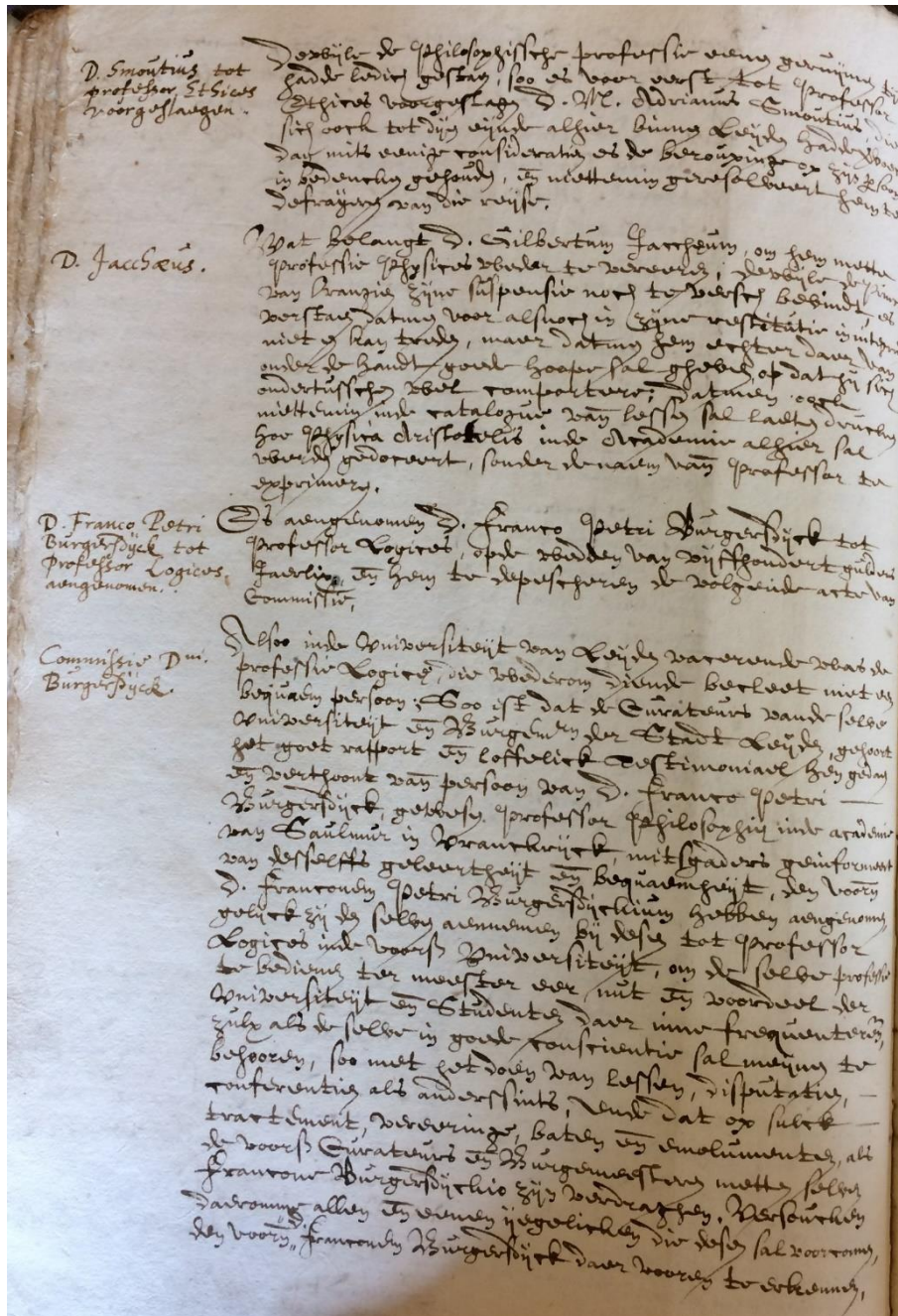
De stukken tekst zijn op de leeszaal Bijzondere Collecties van de UB Leiden gefotografeerd en vervolgens getranscribeerd. In de transcripties zijn opsommingen van namen, aantekeningen in de kantlijn en het kopje dat aan de notulen van elke vergadering voorafgaat en

¹⁵ In verband met de leesbaarheid zijn de tags uit de transcripties weggelaten. Opgeloste afkortingen, die normaal gesproken met de tag <abb>tekst</abb> zouden worden weergegeven, en ambigue lezingen zijn hier in cursief gezet.

waarin onder andere de aan- en afwezigen worden vermeld, niet opgenomen.

Economie

Voor de eerste drie eeuwen die het corpus beslaat, de zestiende tot en met de achttiende eeuw, zijn documenten die betrekking hebben op de textielnijverheid als bron voor het Economische domein gebruikt, aangezien die de belangrijkste economische activiteit vormde in Leiden gedurende deze eeuwen. Posthumus (1910–1922) heeft in enkele publicaties een vrij uitputtend overzicht gegeven van alle bronnen tussen 1333 en 1795 die met de Leidse textielnijverheid te maken hebben. Dit zijn met name ordonnanties en resoluties vanuit de stedelijke overheid en rekesten vanuit de lakenindustrie aan Gerecht en/of burgemeesters. Het materiaal uit de eerste drie eeuwen van het *LOL*-corpus is dan ook geselecteerd uit de bronnen van Posthumus. Posthumus heeft zelf transcripties gemaakt van alle bronnen, maar deze zijn niet diplomatisch (ze volgen onder andere niet de regelafbrekingen van het originele document en zijn aangepast aan de huidige spellingconventies) en slaan soms een deel van de tekst over. Daarom zijn bij de geselecteerde transcripties van Posthumus ook de originele documenten gezocht, die te vinden zijn in het archief Erfgoed Leiden en Omstreken. Dit was geen makkelijke opgave, aangezien Posthumus geen inventarisnummers geeft bij zijn bronnen maar alleen naar bepaalde boeken verwijst (zoals het Gerechts- of Burgemeestersdagboek) met een folionummer en de datum waarop het betreffende stuk tekst is geschreven; voor losbladige documenten volstaat hij met een verwijzing als ‘los document in de portefeuille van de Lakenhal’. Soms konden deze documenten toch nog teruggevonden worden met enkele zoekacties in de online inventarisbeschrijvingen van het archief en dan zijn deze ook aangevraagd en gefotografeerd, maar als dit niet lukte, is het betreffende stuk terzijde geschoven en vervangen door een ander document dat wel op te sporen was in het archief. Vervolgens zijn met behulp van foto's van de originele bronnen de transcripties van Posthumus aangepast aan de transcriptieconventies van het *LOL*-corpus (zie 4.3.2). (Lees verder op pagina 70.)



Afbeelding 4.1. Fragment van de notulen van de vergadering van curatoren en burgemeesters van de Universiteit Leiden, 23 maart 1620. Collectie UB Leiden AC1 21.

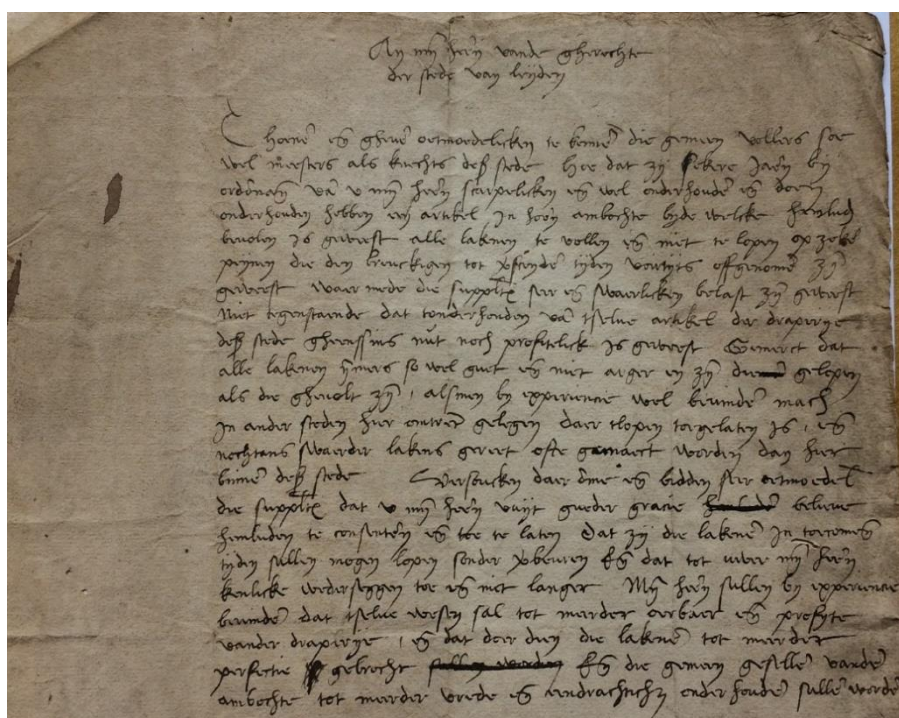
Dewyle de philosophissche professie eenen geruymen tyt hadde ledich gestaen, soo es voor eerst tot professor Ethices voorgeslagen D.M. Adrianus Smoutius, die sich oock tot dyen eynde alhier binnen Leyden hadde *vervoecht*, Dan mits eenige consideratien es de beroupinge op zyn *persoon* in bedencken gehouden, *ende* niettemin geresolveert hem te Defrayeren van Die reyse.

Wat belangt D. Gilbertum Jaccheum, om hem mette professie physices weder te vereeren, dewyle de prince van Orangien zyne suspensie noch te versch bevindt, es verstaen datmen voor alsnoch in zyne restitutie in *integren* niet en kan treden, maer datmen hem echter daer van onder de handt goede hoope sal gheven, op dat hy sich ondertusschen wel comportere: Datmen oock niettemin inde catalogus van~ lessen sal laeten drucken hoe physica Aristotelis inde Academie alhier sal werden gedoceert, sonder de naem van~ professor te exprimeren.

Es aengenomen D. franco petri Burgersdyck tot professor Logices, opde wedden van vyffhondert guldens Jaerlix, *ende* hem te depescheren de volgende acte van Commissie,

Alsoo inde vniversiteyt van Leyden vacerende was de professie Logicae, die wederom Diende becleet met een bequaem persoon; Soo ist dat de Curateurs vande selve vniversiteyt *ende* Burgem~ren der Stadt Leyden, behoort het goet rapport *ende* loffelick Testimoniael hen gedaen *ende* verthoont van~ persoon van D. franco petri Burgersdyck, gewesen professor philosophien inde academie van Saulniur in vranckryck, mitsgaders geïnformeert van desselffs geleertheyt *ende* bequaemhey, den voorn~ D. franconen petri Burgersdyckium hebben aengenomen, gelyck zy den selven aennemen by desen tot professor Logices inde voors~ vniversiteyt, om de selve professie te bedienen ter meester eer nut *ende* voordeel der vniversiteyt *ende* Studenten Daer inne frequenterende, zulx als de selve in goede conscientie sal meynen te behooren, soo met het Doen van Lessen, Disputatien, conferentien als anderssints, ende dat op sulcx tractement, vereeringe, baten *ende* emolumenten, als de voors~ Curateurs *ende* Burgermeesteren metten selven francone Burgersdyckio zyn verdraghen. versoucken Daeromme allen *ende* eenen yegelicken die desen sal voorcomen, den voorn~ D. franconem Burgersdyck Daer vooren te erkennen,

Voor de negentiende eeuw is een andere procedure gevolgd dan voor de zestiende tot en met de achttiende eeuw. De textielnijverheid was aan het einde van de achttiende eeuw vrijwel verdwenen uit Leiden (vandaar dat het overzicht van Posthumus in 1795 stopt), waardoor voor de laatste eeuw van het corpus niet uit Posthumus (1910–1922) geput kon worden. De dataverzameling voor deze periode ging dan ook op een andere manier in zijn werk. Allereerst heb ik de digitale inventarissen van relevante archieven in het Erfgoed Leiden en Omstreken (Leidse bedrijven als de Leidse Zoutkeet, het Stadsarchief en de archieven van de textielhallen en de gilden, die voor de eerste twee decennia van de negentiende eeuw nog wel relevant materiaal bleken te bevatten) doorzocht op het voorkomen van rekestten, ordonnanties en/of resoluties die iets te maken hadden met het economische leven in Leiden. Daarnaast heb ik secundaire literatuur over de Leidse textielnijverheid geraadpleegd, waar ook enkele verwijzingen naar dit soort bronnen in stonden.



Afbeelding 4.2. Verzoek van de vollers van Leiden aan het Gerecht te Leiden, 1535. Collectie Erfgoed Leiden en Omstreken, archief van de Gilden, nummer 1283.

Een probleem was dat er vanwege een Stadhuisbrand in 1929 niet veel over is van de communicatie tussen het stadsbestuur en de lokale bedrijven. Desondanks kon er uiteindelijk genoeg materiaal opgespoord worden. Het materiaal is met name in de tweede helft van de negentiende eeuw door de schaarsheid soms echter wel iets anders van aard dan in de vroegere perioden. Hoewel het nog steeds om besluiten en verzoeken gaat, gaat niet alle correspondentie tussen het stadsbestuur en Leidse bedrijven: enkele verzoeken zijn aan een andere overheidsinstantie (zoals de Provincie) gericht en sommige documenten die zijn geselecteerd betreffen kladbrieven van de stadsarchitect aan Leidse bedrijven, waarmee de eerste reageert op een verzoek tot (ver)bouw van het betreffende Leids bedrijf.

In principe zijn hele documenten getranscribeerd; alleen ondertekeningen en opsommingen van namen zijn niet opgenomen in de transcripties.

An my~ heeren vande gherechte
der stede van leyden

Thoenen *ende* gheuen oetmoedelicken te kennen die gemeen vollers soe wel meesters als knechts *deser* stede hoe dat zy sekere Jaeren by ordo~nan~ va~ v my~ heeren scarpelicken *ende* wel onderhouden *ende* doen onderhouden hebben een artikel In hoeren ambochte byde welcke hemluden beuolen Is geweest alle lakenen te vollen *ende* niet te lopen op zek~ peynen die den breuckigen tot *versceyden* tyden voirtyts offgenomen zy~ geweest Waer mede die suppli~ seer *ende* swaerlicken belast zy~ geweest niet tegenstaende dat tonderhouden va~ tselue artikel der draperye *deser* stede gheenssins nut noch profitelick is geweest Gemerct dat alle lakenen y~mers so wel guet *ende* niet arger en zyn die gelopen als die gheuolt zy~, als men by experientie wel beuinden mach In ander steden hier omtrent gelegen daer tlopen toegelaten js, *ende* nochtans swaerder lakens gereet ofte gemaect Worden dan hier binnen *deser* stede versoucken daer o~me *ende* bidden seer oetmoedel~ die suppli~ dat v my~ heeren vuyt gueder gracie belieue hemluden te consenteren *ende* toe te laten Dat zy die lakenen jn toecomende tyden sullen mogen lopen sonder *verbeuren* *ende* dat tot uwer my~ heeren kenlicke wederseggen toe *ende* niet langer My~ heeren sullen by experientie beuinden dat tselue wesen sal tot meerder oerbaer *ende* profyte vander draperye *ende* dat doer dien die lakenen tot meerder perfectie gebrocht *ende* die gemeen gesellen vanden ambochte tot meerder vrede *ende* eendrachtich~ onderhouden sullen worden

Liefdadigheid

Erfgoed Leiden en Omstreken herbergt veel archieven van charitatieve organisaties: van de meeste hofjes, de gasthuizen en het Heilige Geestweeshuis zijn uitgebreide inventarissen aanwezig. Een genre dat veel voorkomt in deze archieven, zijn testamenten waarin (een deel van) een erfenis wordt geschonken aan een of meer van deze liefdadigheidsinstellingen. Het voordeel van deze testamenten is dat die gedurende de gehele periode die het *LOL*-corpus beslaat in de archieven te vinden zijn en in de meeste perioden in grote mate aanwezig zijn. Om die reden is ervoor gekozen om testamenten met legaten aan Leidse liefdadigheidsinstellingen als het genre te selecteren dat het domein Liefdadigheid in het *LOL*-corpus moet representeren.

De testamenten met schenkingen waren te vinden in alle archieven van liefdadigheidsinstellingen in de inventaris Erfgoed Leiden en Omstreken: alle hofjes, gasthuizen en weeshuizen, maar ook in de archieven van kerken en kloosters. Alleen testamenten die opgesteld zijn door notarissen die in Leiden resideerden en van erflaters die zelf in Leiden woonden zijn opgenomen in het corpus. Daarnaast is er ook gecontroleerd of in de testamenten daadwerkelijk melding werd gemaakt van een schenking aan een Leidse liefdadigheidsinstelling. De meeste testamenten zijn extracten waarin alleen de schenking aan de liefdadigheidsinstelling wordt beschreven, maar er zijn ook enkele hele testamenten bij en ook enkele testamenten waarin de stichting van een hofje wordt geregeld. In principe zijn hele extracten en testamenten getranscribeerd, tenzij deze de 1.250 woorden overschreden (het maximumaantal woorden per tekst volgens de criteria van het *LOL*-corpus, zie de toelichting aan het begin van paragraaf 4.3.1), of als het totale woordenaantal van 5.000 woorden voor de betreffende periode al was bereikt. Passages met kostenoverzichten (nodig voor de registratie van het testament) zijn niet opgenomen in de transcripties.

Er zijn ook andere bronnen dan testamenten aanwezig in de archieven van het weeshuis en van de hofjes die voor het *LOL*-corpus geselecteerd hadden kunnen worden. Deze bronnen – weesbrieven en sollicitatiebrieven – zijn zeer interessant voor de *language history from below*-benadering, maar om verschillende redenen zijn die uiteindelijk niet in het *LOL*-corpus opgenomen.

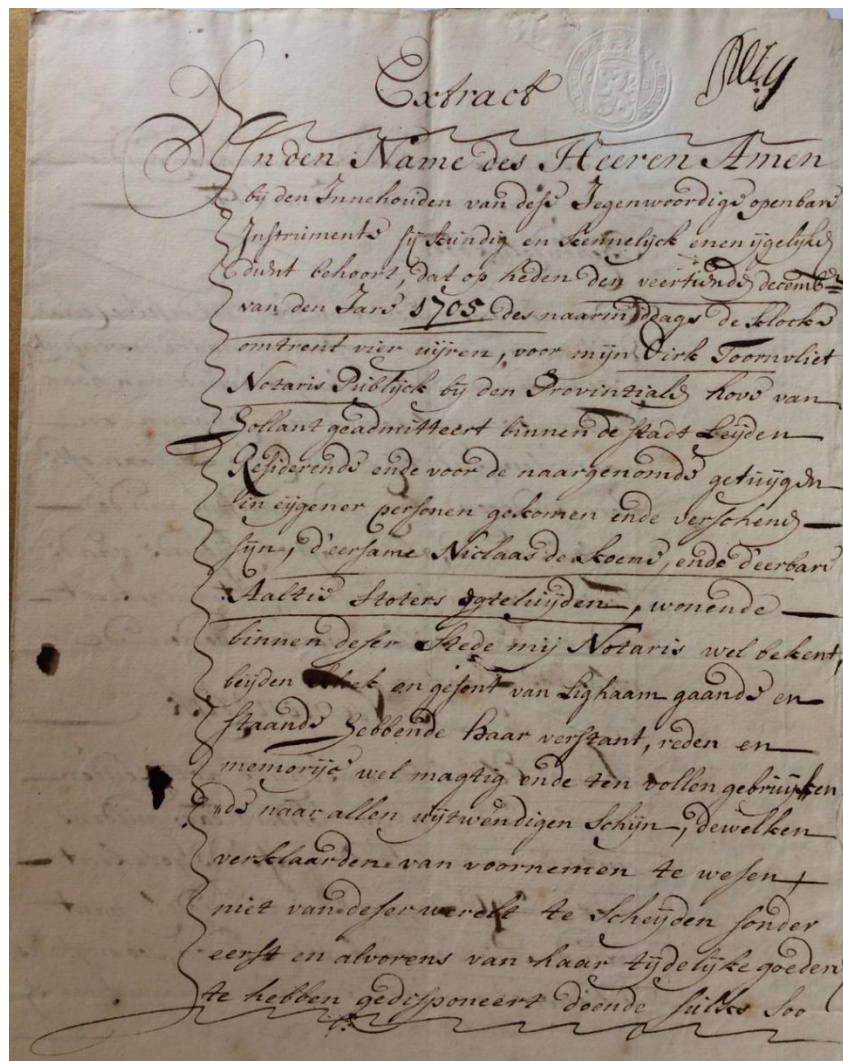
De weesbrieven uit het archief van het weeshuis zijn brieven die door wezen zijn geschreven aan de leiding van het weeshuis: de regenten of de binnenvader en -moeder. Ze zijn geschreven door wezen die er zelf

voor hadden gekozen om het weeshuis te verlaten en door wezen die waren weggestuurd uit het weeshuis wegens wangedrag. Als strafmaatregel moesten zij met een VOC- of WIC-schip mee naar de Oost of de West, of naar een veenkolonie in Drenthe (Van der Wiel 2010: 88–91). In de brieven betuigen de wezen spijt voor hun wandaden, vragen ze de regenten of zij terug mogen komen naar het weeshuis, bedanken ze de bestuurders voor hun goede zorg of verzoeken ze om bepaalde goederen te ontvangen. Voor de historische sociolinguïstiek is dit zeer interessant materiaal, omdat de wezen deze brieven meestal zelf hebben geschreven. Zoals in hoofdstuk 3 (paragraaf 3.4) naar voren kwam, waren veel van de wezen die in het weeshuis terechtkwamen afkomstig uit de armere lagen van de bevolking. Hun brieven maken het mogelijk het taalgebruik van een van de ‘onderbelichte groepen in de samenleving’ (Simons 2013: 20) te bestuderen, iets waar in de historisch-sociolinguïstische *language history from below*-benadering de aandacht op wordt gevestigd.

Ook sollicitatiebrieven uit de archieven van de hofjes, waarin personen verzoeken om een huisje in het betreffende hofje te betrekken, zijn interessant materiaal voor de *language history from below*-benadering. Net als bij de weesbrieven moeten de afzenders van de sollicitatiebrieven uit de lagere klassen van de bevolking afkomstig zijn geweest, aangezien de hofjes voor arme personen bestemd waren.

Het probleem met zowel de weesbrieven als de sollicitatiebrieven is echter dat er voor beide genres maar voor een beperkte periode materiaal te vinden is: de weesbrieven die in het archief aanwezig zijn, zijn afkomstig uit de tweede helft van de achttiende en de eerste helft van de negentiende eeuw en de sollicitatiebrieven komen uit de negentiende eeuw. De voorkeur ging bij alle domeinen echter uit naar een genre waarin zo veel mogelijk perioden gerepresenteerd zijn. Bovendien is het bij zowel de weesbrieven als bij de sollicitatiebrieven lastig om te bepalen of de brieven autograaf zijn of niet, dat wil zeggen: of de brieven geschreven zijn door de afzenders zelf, of dat iemand anders de brief voor hen heeft geschreven. Met name bij de sollicitatiebrieven bleek het zeer lastig om voldoende brieven te vinden die onmiskenbaar door de afzender waren geschreven: in sommige brieven werd bijvoorbeeld expliciet gezegd dat een bepaalde brief namens iemand anders was geschreven of werd er niet in de *ik/wij*-vorm, maar in de *hij/zij*-vorm geschreven. Deze brieven konden niet opgenomen worden in het *LOL*-corpus en aangezien er vrij veel brieven waren waarbij het twijfelachtig was of het wel een autografe brief was, bleef er maar weinig materiaal over voor het corpus. Vanwege de schaarsheid van

het materiaal en het beperkte aantal perioden waarin dit materiaal aanwezig is, is er dus voor gekozen om niet de wees- en sollicitatiebrieven, maar de testamenten voor het domein Liefdadigheid te kiezen.



Afbeelding 4.3. Fragment van een extract van een testament, 25 januari 1707. Collectie Erfgoed Leiden en Omstreken, archief van de Gasthuizen, nummer 1170.

Extract

In den Name des Heeren Amen
 by den Innehouden van dese Jegenvoordige openbare
 Instrumente sy kundig en kennelyck enen ygelyken
 dient behoort, dat op heden den veertienden december
 van den Jare 1705 des naarmiddags de klokke
 omtrent vier uyren, voor myn Dirk Toornvliet
 Notaris Publyck by den Provintialen hove van
 hollant geadmitteert binnen de stadt Leyden
 Residerende ende voor de naargenomde getuygen
 in eygener personen gekomen ende verschenen
 syn, d'eersame Niclaas de koene, ende d'eerbare
 Aaltie stoters egteluyden, wonende
 binnen deser stede my Notaris wel bekent,
 beyden kloek en gesont van Lighaam gaande en
 staande hebbende haar verstant, reden en
 memorye wel magtig ende ten vollen gebruyken
 de naar allen uytwendigen schyn, dewelken
 verklaarden van voornemen te wesen,
 niet van deser werelt te scheyden sonder
 eerst en alvorens van haar tydelyke goederen
 te hebben gedisponeert doende sulks soo

Literatuur

Voor het domein Literatuur is gekozen voor het literaire genre toneelstukken. Vanwege de grote populariteit van dit genre bij de rederijderskamers en ook in de eeuwen erna was het te verwachten dat er veel toneelstukken te vinden zouden zijn. Daarnaast wordt het zoeken naar toneelstukken tot 1803 vergemakkelijkt door het bestaan van de Ceneton (Census Nederlands Toneel), waarin de gegevens van bijna 13.000 Nederlandse toneelstukken opgenomen zijn.¹⁶ De meeste toneel-

¹⁶ De Ceneton is beschikbaar via <https://www.let.leidenuniv.nl/Dutch/Ceneton/> (voor het laatst geraadpleegd op 30 augustus 2022).

stukken zijn dan ook met behulp van deze lijst opgespoord. Daarnaast was de DBNL een onmisbare bron voor het vinden van toneelschrijvers en hun toneelstukken.

Het belangrijkste criterium bij de selectie van de toneelstukken was dat de toneelschrijver van het betreffende stuk een sterke band met Leiden moest hebben. Van de meeste beschikbare toneelstukken is de auteur bekend en van veel van hen zijn ook wat bibliografische gegevens bekend in de DBNL en soms ook elders op internet, bijvoorbeeld de geboorte- en/of sterfplaats of het beroep en de plaats waar dit beroep werd uitgeoefend. Op basis van deze gegevens kon bepaald worden hoe sterk de band van een specifieke toneelschrijver met de stad Leiden was. In de ideale situatie zouden alleen toneelteksten geselecteerd zijn waarvan de auteur in Leiden is geboren en gestorven en zijn hele leven in Leiden heeft gewoond. Dit bleek echter een onmogelijke opgave: er konden slechts 7 auteurs gevonden worden voor wie Leiden zowel de geboorte- als sterfplaats was. Voor de selectie van de overige 17 schrijvers van wie uiteindelijk teksten in het domein Literatuur zijn opgenomen, is een andere methode gehanteerd: er is op basis van de specifieke levensloop van de individuele auteur en de context waarin het toneelstuk is geschreven bepaald of de tekst in kwestie als Leids aangemerkt zou worden of niet. Hierbij zijn geen harde criteria gebruikt: ik heb bijvoorbeeld geen minimumaantal jaren gehanteerd waarin de toneelschrijver in Leiden moet hebben gewoond en ook is het doorbrengen van de jeugd in Leiden geen voorwaarde geweest om een schrijver voor het corpus te selecteren. Het criterium dat een auteur aangemerkt moet kunnen worden als 'Leids' om voor het *LOL*-corpus te worden geselecteerd, betekent dus niet dat deze persoon Leidse kenmerken in zijn taalgebruik moet hebben. Ik vat 'Leids' in mijn onderzoek niet op in dialectologisch opzicht: het gaat erom dat het tekstmateriaal dat geselecteerd wordt voor het corpus in een bepaalde periode representatief is voor een sociaal domein in Leiden; het is niet bedoeld als een verzameling teksten waarin Leidse dialectkenmerken voorkomen.

De argumenten op grond waarvan ik een toneelauteur alsnog als Leids heb aangemerkt, zijn heel verschillend. Hieronder volgen een aantal voorbeelden:

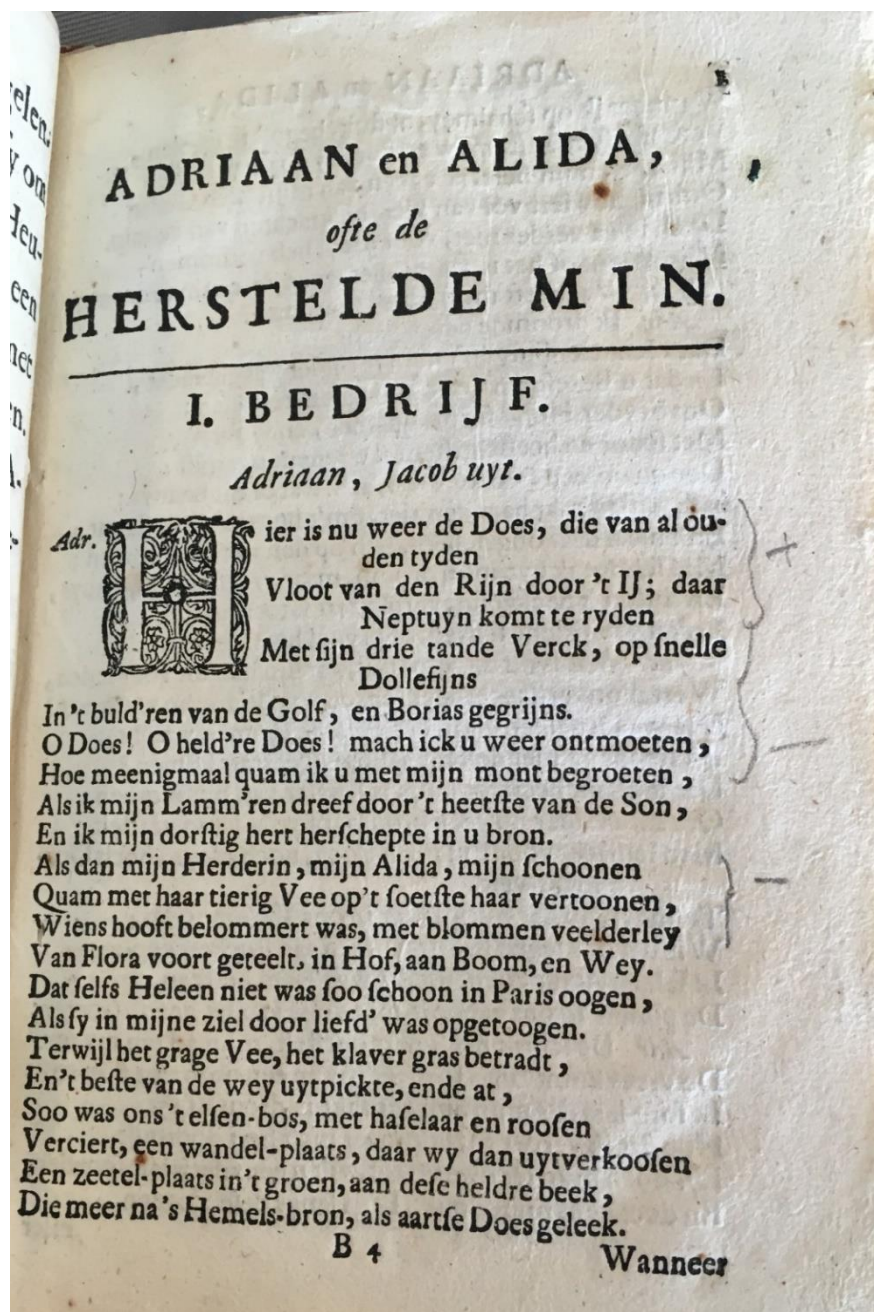
- Niet alleen Leiden zelf, maar ook de directe omgeving van Leiden is als 'Leiden' meegeteld. Zo is van een van de geselecte-

teerde toneelschrijvers, Willem van Heugelenburg, bekend dat deze uit Leiderdorp kwam;

- Enkele auteurs hebben slechts een korte periode in Leiden gewoond (maximaal tien jaar), bijvoorbeeld omdat ze er studeerden. De toneelstukken van hen die in het corpus zijn opgenomen, zijn echter geschreven in de tijd dat zij in Leiden woonden, waardoor de toneelstukken dus in zekere zin producten van Leidse bodem zijn. Dit geldt voor het toneelstuk *Meleager en Atalante* van Lukas Schermer (periode 1700–1749) en *De mislukte smokkelpartij* van Aart Veder (periode 1800–1849);
- Van Jacob van Rijndorp is niet zeker of hij in Leiden heeft gewoond, maar aangezien hij de oprichter en directeur van de Leidse Schouwburg was en het toneelstuk speciaal voor deze schouwburg was geschreven, is ook dit toneelstuk in het corpus opgenomen;
- Van sommige auteurs is de geboorte- en sterfplaats niet bekend of is Leiden geen van deze plaatsen, maar soms is wel uit andere bibliografische informatie op te maken dat ze in elk geval een tijd in Leiden hebben gewoond. Frans van Aken (periode 1750–1799) is bijvoorbeeld geboren en gestorven in Amsterdam, maar hij zou wel lange tijd in Leiden hebben gewoond en daar als horlogemaker hebben gewerkt. Dit geldt bijvoorbeeld ook voor Willem Bilderdijk (periode 1800–1849) en voor Johannes Karel de Regt (periode 1850–1899).

Bij de selectie van het materiaal moest het begrip ‘Leids’ noodzakelijkerwijs steeds breder worden opgevat naarmate er minder materiaal voor een periode beschikbaar was. In elk geval kan van alle auteurs van wie teksten zijn opgenomen in het corpus worden gezegd dat deze in elk geval een tijd in Leiden hebben gewoond. Wel heb ik altijd de voorkeur gegeven aan personen die hun hele leven of bijna hun hele leven in Leiden hebben gewoond; alleen als er hier niet genoeg van waren, kwamen er andere argumenten in beeld om een bepaald toneelstuk te selecteren voor het corpus.

Bij de selectie van toneelstukken is er naast de band van de auteur met Leiden ook rekening mee gehouden dat alleen originele teksten van de betreffende schrijver werden gebruikt, en bijvoorbeeld geen bewerkingen of vertalingen van toneelstukken.



Afbeelding 4.4. Fragment uit het toneelstuk *Adriaan en Alida* door Willem van Heugelenburg, 1689. Collectie UB Leiden 1092 E 65.

ADRIAAN en ALIDA,
 ofte de
 HERSTELDE MIN.
 I. BEDRIJF.
 Adriaan, Jacob uyt.
 Adr. Hier is nu weer de Does, die van al ou-
 den tyden
 Vloot van den Rijn door 't IJ; daar
 Neptuyn komt te ryden
 Met sijn drie tande Verck, op snelle
 Dollefijns
 In't buld'ren van de Golf, en Borias gegrijns.
 O Does! O held're Does! mach ick u weer ontmoeten,
 Hoe meenigmaal quam ik u met mijn mont begroeten,
 Als ik mijn Lamm'ren dreef door 't heetste van de Son,
 En ik mijn dorstig hert herschepte in u bron.
 Als dan mijn Herderin, mijn Alida, mijn schoonen
 Quam met haar tierig Vee op't soetste haar vertoonen,
 Wiens hoofd belommert was, met blommen veelderley
 Van Flora voort geteelt, in Hof, aan Boom, en Wey.
 Dat selfs Heleen niet was soo schoon in Paris oogen,
 Als sy in mijne ziel door liefd' was opgetoogen.
 Terwijl het grage Vee, het klaver gras betradt,
 En't beste van de wey uytpickte, ende at,
 Soo was ons 't elsen-bos, met haselaar en roosen
 Verciert, een wandel-plaats, daar wy dan uytverkoosen
 Een zeetel-plaats in 't groen, aan dese heldre beek,
 Die meer na 's Hemels-bron, als aartse Does geleet.

Wanneer

De meeste toneelstukken zijn als boekjes uitgegeven, die ofwel in de UB in Leiden ter inzage liggen bij de afdeling Bijzondere Collecties, ofwel via Google Books geraadpleegd kunnen worden. Van al deze toneelstukken is de eerste druk gebruikt voor de transcriptie. Twee toneelstukken uit de periode 1550–1599 waren alleen beschikbaar in handschrift, die in Erfgoed Leiden en Omstreken respectievelijk bij de Koninklijke Bibliotheek in Den Haag opgevraagd konden worden. Voor een van de handschriften, *Loterijspel* van Jan van Hout, kon ook gebruikgemaakt worden van een reeds beschikbare transcriptie (Koppenol 1998). Deze transcriptie is vergeleken met het originele handschrift en aangepast aan de transcriptieconventies van het *LOL*-corpus.

Op enkele punten moest afgeweken worden van de principes van het corpus door gebrek aan materiaal. Voor de eerste helft van de

zestiende eeuw is helemaal geen materiaal aangetroffen in de (digitale) archieven. Eén van de gebruikte toneelstukken die in handschrift is overgeleverd (*Wie voirmaels waeren de victoriöste*), toegeschreven aan Cornelis Meesz. van Hout) zou in 1548 zijn geschreven, maar aangezien de datering van het handschrift op 1552 is gesteld, is dit opgenomen bij de toneelstukken van de periode 1550–1599. Daarnaast zijn voor de perioden 1550–1599, 1650–1699 en 1850–1899 slechts drie in plaats van vier toneelstukken gebruikt voor het corpus, waarvan twee toneelstukken in de periode 1850–1899 ook nog dezelfde auteur hadden. Uitgebreide zoektochten in de beschikbare overzichten van toneelstukken en in de secundaire literatuur leverde echter niet meer materiaal op en daarom is ervoor gekozen om voor deze perioden een te transcriberen woorden-aantal van ongeveer 1.700 per toneelstuk aan te houden in plaats van 1.250. Alle toneelstukken uit alle perioden zijn vanaf het eerste bedrijf getranscribeerd tot het maximumaantal woorden voor dat toneelstuk (ofwel 1.250 ofwel 1.700) was bereikt.

Privéleven

Voor het Privéleven is gekozen voor privébrieven. Dit is een genre dat ruim aanwezig is in de archieven van Leidse families. Documenten als dagboeken en reisverslagen zijn veel moeilijker vindbaar, wat dit een minder goede optie maakte.

Er waren al enkele transcripties van privébrieven uit de perioden 1600–1649, 1750–1799 en 1800–1849 beschikbaar die studenten in het kader van een MA-werkgroep historische sociolinguïstiek hadden gemaakt. Deze brieven zitten in de collecties van de UB Leiden en Erfgoed Leiden en Omstreken. Enkele hiervan zijn geselecteerd voor het *LOL*-corpus. Met behulp van foto's van de originele documenten zijn de transcripties nagekeken en aangepast aan de transcriptieconventies van het *LOL*-corpus.

De familiearchieven die te vinden zijn in de collectie van Erfgoed Leiden en Omstreken zijn het uitgangspunt geweest voor de verdere zoektocht naar privébrieven. Die leverde echter voor enkele perioden, en met name de perioden 1650–1699 en 1700–1749, niet de gewenste 5.000 woorden per periode op. Daarom is ook gezocht in de inventarissen van het Stadsarchief Amsterdam en het Nationaal Archief te Den Haag. Hier konden genoeg brieven van Leidenaren worden gevonden om aan de 5.000 woorden te komen. Uiteindelijk voldoet alleen de zestiende eeuw

niet aan het gestelde woordenaantal: voor de periode 1500–1549 zijn helemaal geen brieven gevonden en voor de periode 1550–1599 komt het woordenaantal slechts op 4.449 uit. Voor deze perioden is het echter zeer de vraag of er (meer) brieven van Leidenaren te vinden zullen zijn als het zoekgebied wordt uitgebreid naar nog weer andere archieven, aangezien het overgeleverde brievenmateriaal voor de zestiende eeuw sowieso schaars is.

Bij de selectie van de brieven is op een aantal dingen gelet. Allereerst was het, net als bij de toneelauteurs voor het domein Literatuur, van belang dat de briefschrijvers een duidelijk aantoonbare band hadden met Leiden. Dit was vaak redelijk gemakkelijk te bepalen, aangezien van de meeste brieven bekend is wie ze geschreven hebben en de meeste schrijvers uit vooraanstaande families kwamen, waarvan veel informatie op het internet te vinden is. Vaak konden dan ook de stamboom, de geboorte- en sterfdatum en -plaats en meestal een levensbeschrijving achterhaald worden. Brieven van personen die Leiden als geboorte- en sterfplaats hadden, zijn zonder meer opgenomen in het corpus. Van personen die ofwel in Leiden waren geboren ofwel gestorven, die Leiden niet als hun geboorteplaats en ook niet als hun sterfplaats hadden of van wie de geboorte- en sterfplaats helemaal niet bekend waren, is net als bij de toneelschrijvers in het domein Literatuur verdere informatie opgezocht over de levensloop van deze personen. De meeste van deze personen kwamen uit vooraanstaande Leidse families, zoals de familie Hubrecht, Dozy, Luzac of De la Court, en waren in de loop van hun leven ergens anders gaan wonen. Ook hun brieven zijn als Leids aangemerkt, maar er gold wel: hoe langer de persoon in kwestie in Leiden had gewoond, hoe eerder deze werd geselecteerd voor het *LOL*-corpus, afhankelijk van de hoeveelheid beschikbare brieven voor een bepaalde periode. Over het algemeen was het voor het Privédomein wel eenvoudiger om materiaal te vinden van auteurs die lang in Leiden hadden gewoond dan voor het domein Literatuur. Veel van de geselecteerde brieven zijn ook in Leiden zelf geschreven. Met name in de zestiende en zeventiende eeuw was het beschikbare brievenmateriaal echter enigszins schaars en was er van sommige personen weinig bibliografische informatie beschikbaar, waardoor het begrip 'Leids' soms wat breder opgevat moest worden. Enkele voorbeelden:

- Voor de periode 1550–1599 kon enkel geput worden uit het archief van de familie Van der Meulen. Deze familie kwam pas in 1591 in Leiden wonen, nadat zij eerst in Antwerpen en Bremen

hadden gewoond. Er is daarom alleen correspondentie van na 1591 uit hun archief in het corpus opgenomen. Het archief bevatte ook enkele brieven van familieleden van de familie Van der Meulen van wie geen bibliografische informatie gevonden kon worden. Eén brief, geschreven door ene Katharina van der Bleke, werd geschreven vanuit Leiden. Omdat deze brief wel in Leiden was geschreven, dit de enige informatie was waar ik mij op kon baseren en voor deze periode het aantal woorden van 5.000 sowieso niet kon worden gehaald, heb ik deze brief ook opgenomen in het corpus.

- Van Adriana Fabricius en Engeltje Fabricius (periode 1650–1699) is de geboorteplaats niet Leiden en hun sterfplaats is onbekend, maar beiden schreven hun brieven onder andere vanuit Leiden en ze zijn ook beiden met Leidenaars getrouwd geweest, dus om deze reden zijn hun brieven ook in het corpus opgenomen.

Zeker voor de tweede helft van de achttiende en de negentiende eeuw was er een ruime hoeveelheid briefmateriaal beschikbaar in de archieven van personen die een sterke band hadden met Leiden, waardoor de keuze voor bepaalde briefschrijvers ook nog van andere criteria kon afhangen. Indien het beschikbare materiaal dat toeliet, gaf ik de voorkeur aan brieven die geschreven waren aan directe familieleden (ouder/kind-, man/vrouw- en broer/zus-relaties hadden bijvoorbeeld de voorkeur boven oom/neef- of neef/nicht-relaties) en streefde ik naar zo veel mogelijk verschillende briefschrijvers en een gelijk woordenaantal van mannelijke en vrouwelijke schrijvers per periode. Ook gaf ik, indien mogelijk, de voorkeur aan een ruime verspreiding van de brieven over de gehele periode van vijftig jaar. Er is ook voor alle perioden op gelet dat de brieven autograaf waren.

Uiteindelijk heeft alleen de zestiende eeuw problemen opgeleverd bij de opbouw van het corpus. Voor de eerste helft van de zestiende eeuw is helemaal geen privécorrespondentie gevonden van Leidenaren en in de tweede helft van de zestiende eeuw waren er, zoals hierboven al opgemerkt, te weinig brieven om aan het totale woordenaantal van 5.000 te voldoen. Daarnaast zijn er van een briefschrijfster, Hester de la Faille, 2.190 woorden in het corpus opgenomen voor de tweede helft van de zestiende eeuw. Dit is bijna het dubbele van de 1.250 woorden die volgens de principes van het corpus eigenlijk per schrijver per periode mochten worden opgenomen. Aangezien er echter geen ander materiaal

beschikbaar was voor deze periode, waren er twee opties: of er werd voor gekozen om de extra woorden van Hester de la Faille op te nemen zodat het woordenaantal voor deze periode de 5.000 zo dicht mogelijk zou benaderen, of er werd strikt vastgehouden aan de 1.250 woorden per briefschrijver, wat betekent dat het woordenaantal voor de tweede helft van de zestiende eeuw bijna 1.000 woorden lager zou uitvallen. Ik heb de voorkeur gegeven aan een zo groot mogelijk woordenaantal, zodat er genoeg woorden beschikbaar zouden zijn om de casestudies mee uit te voeren.

Voor het corpus zijn alleen hele brieven getranscribeerd, indien beschikbaar met de adressering erbij. Voor de meeste brieven afkomstig uit de archieven van de familie Van der Meulen die voor het *LOL*-corpus waren geselecteerd (voor de perioden 1550–1599 en 1600–1649), was al een transcriptie beschikbaar in Hendriks (1998). Van enkele andere brieven uit de periode 1600–1649 waren ook al transcripties bij de archiefstukken gevoegd. Deze bestaande transcripties zijn gecontroleerd aan de hand van het origineel en aangepast aan de transcriptieconventies van het *LOL*-corpus.

Publieke Opinie

De krant is het genre dat is gekozen voor het domein Publieke Opinie. De *Leydse(/Leidse/Leidsche) Courant* bestond vanaf 1686 en was lange tijd de enige Nederlandse krant in Leiden. Dit is dan ook de enige bron geweest voor het domein Publieke Opinie. De meeste krantenedities zijn in gescande vorm te raadplegen via de website van Delpher en via Erfgoed Leiden en Omstreken. Net als bij het Privédomein waren ook deze transcripties deels al door MA-studenten gemaakt. Deze zijn echter nog wel nagekeken en daarbij zijn stukken geschrapt waarin veel namen werden genoemd. Ook advertenties (geboortes, overlijdens of reclames) zijn uit de transcripties gehaald en vervangen door nieuwsberichten. Aangezien de krant pas vanaf 1686 verscheen, is er voor de eerste drie perioden van het *LOL*-corpus geen materiaal beschikbaar. Bij de rest van de perioden zijn telkens vier of meer kranten uitgekozen uit het midden van de betreffende periode waaruit per krant maximaal 1.250 woorden werden getranscribeerd.

26 Febr. 1827

Zoo is het dan, Leve Loo! al van de Meisjes oort of pret
 by u geweest; het is mogelijk dat het deff Waarden
 buij was die u amuseerde, indien het my mees
 bekend was geweest, zoud. ik het hebben kunnen wagen,
 want zynde heeft in de vorige week ons een Thee-
 Visite gegeven; en moeten my deswegens in een
 atelinge onschuld tot naer inlichting bekomen te
 hebben blyven; binnen den dag die gy by Tante
 Angelis hebt doorgebracht, hebt ge immers de andere
 niet met slapen doorgebracht, want van een ander
 bedryf met gy niets. Of een wandelogen-
 schap in zult gy wel welde kousen aantrekken
 hebben, en de spate bewaard tot het eens regenen
 magt, om dat het dan min of meer ruil of slibberig
 op de steenen straten zoud. kunnen zijn; de
 goede welstand van alle de vrienden te vernemen, was
 my zeer aangenaam, en ware het anders zoud. het
 my bestelyk leed zijn, doed aan alle myn lieve
 presentatie; ik weet nog niet hoe of het met
 jehommaken, grote West, Troven een sal uitkomen,
 om Charles te hebben, dan moet het nader bepaald
 worden; ik heb my by de loep geuniformeerd, of by
 geene gelotintBriem, meender med. neend, dan nog
 geen berge deessons; ik weet niet, dat hier iets
 is omgevallen, "Dine altoude waardig; alle zijn my
 my nill; Straks viel'er door een stukje harden
 sneeuw en Reut en van de Lantaarn waar door my
 bespeeld hebben, dat die rood bloed heeft, als ook
 Maatje, Zynde beiden een lekkersje gekroet, doch
 ontbreekt u over de bronnen niet; vanmorgen vernam
 ik dat Mevrouw Salomon gestorven is Overleden, ik
 met niet van hare ziekte, wel dat sy van tyd tot
 tyd sukkelde, zeg dat niet schuchlyk aan de goede
 Vrienden Breyers, want misschien heeft haar de
 Overleden min of meer byzonder gekend; Lee en
 Cabot sluiten hunne groeten hier by in, van Frans,

Afbeelding 4.5. Fragment uit een brief van Jan Pieter van Klinkenberg Dozy aan zijn dochter, 26 februari 1827. Collectie Erfgoed Leiden en Omstreken, archief Familie Dozy, nummer 459.

zoo is het dan lieve Lot! al van de Mettepoort af pret
 by u geweest, het is mogelyk dat het de Waarden
 burg was die u amuseerde, indien het my vroeger
 bekend was geweest, zoude ik het hebben kunnen vragen,
 want zynEd. heeft in de vorige week ons eene Thee-
 visite gegeven; nu moeten wy deswegens in eene
 akelige onzekerheid tot nader inligting bekomen te
 hebben blyven: buiten den dag die gy by Tante
Acnelis hebt doorgebracht, hebt ge immers de andere
 niet met Slapen doorgebracht, want van uwe andere
 bedryven meld gy niets. op uwe wandelingen
 vertrouwde ik zult gy wel nette koussen aangetrokken
 hebben, en de Swarte bewaard tot het eens regenen
 mogt, om dat het dan min of meer vuil of slibberig
 op 's Heeren Straaten zoude kunnen zyn: de
 goede welstand van alle de Vrienden te vernemen, was
 my zeer aangenaam, en ware het anders zoude het
 my hertelyk leed zyn, *doed* aan alle myne dienst
 presentatie: ik weete nog niet, hoe of het met
 schoonmaken, grote Wast, Turven enz. zal uitkomen,
 om Charles te hebben, dan moet dit nader bepaald
 worden: ik heb my by de looper geïnformeerd, of hy
 geen gesloten Brieven meerder medeneemd, dan nog
 geen berigt deswegens: ik weete niet, dat hier iets
 is voorgevallen, uwe attentie waardig: alle zyn wy
 vry wel: straks viel 'er door een stukje harde
 sneeuw een Ruit in van de Lantaarn waar door wy
 bespeurd hebben, dat Cie rood bloed heeft, als ook
 naatje, zynde beiden een sikkepitje gekwetst, doch
 ontrust u over die wonden niet: vanmorgen vernam
 ik dat Mevrouw Solomon gisteren is overleden, ik
 wist niet van haare Ziekte, wel dat zy van tyd tot
 tyd sukkelde, zeg dit niet schielyk aan de goede
 vrienden Bruyns, want misschien heeft haar E. de
 overledene min of meer byzonder gekend: Cie en
 Cabet sluiten hunne groeten hier by in, van Frans



Afbeelding 4.6. Fragment uit de *Leydse Courant* van 5 januari 1780. Online collectie Delpher.nl.

[Alleen van de linker kolom van afbeelding 4.6 wordt hieronder de transcriptie gegeven.]

RUSLAND.

PETERSBURG den 14 December. De Heer Hablizi, Correspondent van onze Keyzerlyke Academie, heeft in een Brief uit Astracan, gedateerd 24 October, eene merkwaardige Byzonderheid gemeld, hier in bestaande, dat de Overstrooming, welke gewoonlyk in de Vloed Wolga plaats heeft, dit Jaar veel hooger is geweest van in 20 Jaaren Tyds waargenoomen is. In 't midden van April begon het Water te wassen; op den 31 Mey was het op het hoogste en omtrent het midden van July weder als naar gewoonte; doch allerbyzonderst is dat men, een Proef genoomen hebbende, om zig, gelyk in Egypten met de Nyl geschied, van deze Overstrooming te bevoordeelen, daar in zeer gelukkig is geslaagt.

GROOT-BRITTANNIEN.

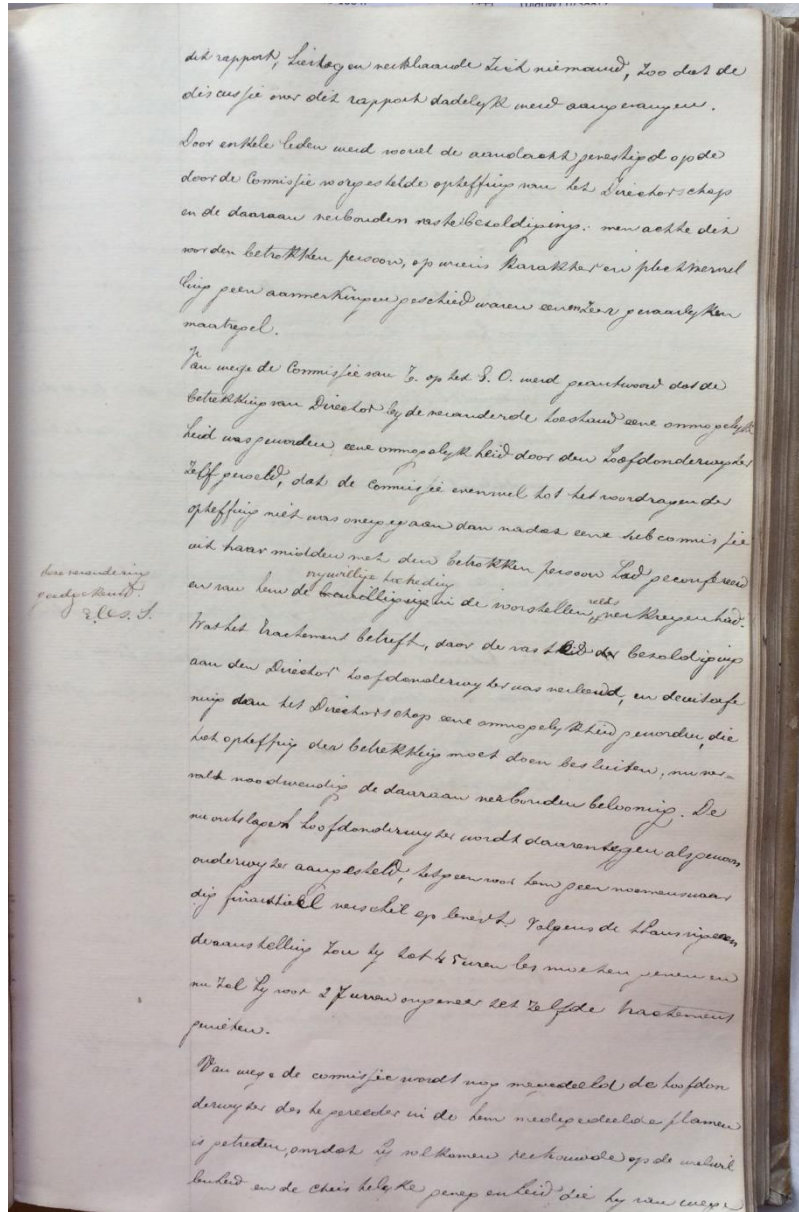
WITHBY den 24 December. Donderdag is by Runswick gestrand en vergaan de Sloep van John Thompson, van Middelburg in Zeeland na Leith bestemd.

PORTSMOUTH den 27 December. Gister is het Oorlogschip the Warwick met negen Retourschepen van onze O.I. Compagnie alhier aangekomen; Daartegen is heden het Esquader, sterk 24 Schepen, daaronder een van 90 St. Kanon, den Admiraal Rodney aan boord hebbende, een van 100, een van 98, en tien van 74 St. Kanon, met alle de na de West-Indiën, Gibraltar enz. bestemde Koopvaarders onder zeil gegaan.

LONDEN den 24 December. Dingsdag lazen de Lords de Bills wegens Ierland voor de eerste keer, welke op dien Dag by de Gemeentens in Comité in order gebragt wierden. Eergister wierden de gem. Bills in 't Hoogerhuis voor de tweede keer gelezen en gister heeft de Koning dezelve goedgekeurd en onder anderen ook een Acie ter herroeping van zekere Acten, in Groot-Brittannien gemaakt, waar by de Koophandel en Commerce van Ierland met uitheemsche Havens bepaald was.

Het Hof heeft genoeg om te verneemen dat de Resolutien van ons Parlement om den Koophandel van Ierland uit te breiden, door het Parlement van dat Koningryk, zo wel als van de gantsche Natie met de grootste Dankbaarheid en Aandoening ontfangen zyn, dus men de gelukkigste Gevolgen te gemoet ziet, en byzonderlyk eene Eendragtigheid van Maatregelen voor het algemeene Welzyn der drie Koningryken.

Voor eenige Dagen is de Gemeene Raad van deze Stad vergaderd geweest ter beraaming van eene gevoeglyke wyze om de Nagedagtenis van wylen den Doorluchtigen Graaf van Chatham te



Afbeelding 4.7. Fragment uit de notulen van de kerkenraadsvergaderingen van de Hervormde Kerk in Leiden, 7 mei 1875. Collectie Leiden Erfgoed en Omstreken, archief Kerkeraad van de Nederlands Hervormde gemeente te Leiden, nummer 17.

dit rapport, hiertegen verklaarde Zich niemand, Zoo dat de discussie over dit rapport dadelyk werd aangevangen. Door enkele leden werd vooral de aandacht gevestigd op de door de Commissie voorgestelde opheffing van het Directorschap en de daaraan verbonden vaste bezoldiging: men achtte dit voor den betrokken persoon, op wiens karakter en *plicht* ling geen aanmerkingen geschied waren eenen Zeer gevaarlyken maatregel.

Van wege de Commissie van t. op het S.O. werd geantwoord dat de betrekking van Director by de veranderde toestand eene onmogelykheid was geworden eene onmogelykheid door den hoofdonderwyzer zelf gevoeld, dat de commissie evenwel tot het voordragen der opheffing niet was overgegaan dan nadat eene subcommissie uit haar midden met den betrokken persoon had geconfereerd en van hem de vrywillige toetreding in de voorstellen reeds verkregen had.

Wat het tractement betreft, daar de vaste bezoldiging aan den Director hoofdonderwyzer was verleend, en de uitoefening van het Directorschap eene onmogelykheid geworden, die tot opheffing der betrekking moet doen besluiten, nu vervald noodwendig de daaraan verbonden belooning. De nu ontslagen hoofdonderwyzer wordt daarentegen als gewoon onderwyzer aangesteld, hetgeen voor hem geen noemenswaardig finantieel verschil op levert. Volgens de thans *ingeven* de aanstelling zou hy tot 45 uren les moeten geven en nu zal hy voor 27 uren ongeveer het zelfde tractement genieten.

Van wege de commissie wordt nog medegedeeld de hoofdonderwyzer des te gereeder in de hem medegedeelde plannen is getreden, omdat hy volkomen vertrouwd op de welwillenheid en de christelyke genegenheid die hy van wege

Religie

Bij de keuze voor een genre voor het domein Religie is bekeken welk soort materiaal het ruimst aanwezig was gedurende de vier eeuwen die het *LOL*-corpus beslaat in de kerkarchieven die in Erfgoed Leiden en Omstreken worden bewaard. Dit bleken de notulen of resoluties van kerkenraden te zijn. Er is materiaal van drie verschillende geloofsstromingen opgenomen in het corpus – hervormd, doopsgezind en evangelisch-luthers – omdat van de kerkenraden van deze drie religies het meeste materiaal, verspreid over een langere periode, aanwezig was. Het feit dat er is gericht op drie protestantse stromingen betekent wel dat er voor de eerste helft van de zestiende eeuw geen materiaal in de archieven aanwezig was, aangezien deze geloofsstromingen toen nog niet bestonden. Voor de eerste helft van de zestiende eeuw hadden er nog

wel teksten van de katholieke kerken toegevoegd kunnen worden aan het corpus, maar ik heb in de archieven geen teksten aangetroffen die vergelijkbaar waren met de teksten uit de andere perioden.

De kerkenraadboeken van de hervormde kerk beslaan de langste periode: van de tweede helft van de zestiende eeuw tot de twintigste eeuw. De doopsgezinden en lutheranen zijn beide vanaf de tweede helft van de zeventiende eeuw tot de twintigste eeuw gerepresenteerd. Omdat een van de principes van het *LOL*-corpus was dat er ten minste vier documenten per domein per periode opgenomen moesten worden in het corpus om de invloed van voorkeuren van individuele schrijvers zo veel mogelijk in te perken, is ervoor gezorgd dat in de perioden waarin alleen de hervormde kerkenraadboeken beschikbaar waren er vier verschillende handen uit die kerkenraadboeken zijn gebruikt, met per hand een totaal aantal woorden van ongeveer 1.250. Voor de perioden waarin het materiaal ook uit de kerkenraadboeken van de doopsgezinden en lutheranen bestond, zijn van elke geloofsstroming stukken van de kerkenraadboeken in ten minste twee verschillende handen gebruikt, wat neerkomt op een totaal van ten minste zes handen voor elk van deze perioden. Stukken van de kerkenraadboeken waarin veel namen voorkwamen (bijvoorbeeld bij de benoeming van diakenen) zijn overgeslagen.

4.3.2 *Transcriptieconventies en transcriptieprocedure*

De documenten die het archiefonderzoek voor de verschillende perioden en domeinen opleverde, zijn alle getranscribeerd aan de hand van door mij opgestelde transcriptieconventies. Voor het opstellen van deze conventies heb ik gebruikgemaakt van de transcriptieconventies van het *Brieven als Buit*-corpus (Nobels 2013: 295–296; Simons 2013: 337–338) en van het *Going Dutch*-corpus (Krogull 2018: 327–329). In navolging van Krogull (2018) heb ik aparte transcriptieconventies gemaakt voor handgeschreven en gedrukte documenten, aangezien beide soorten bronnen in het *LOL*-corpus terecht zijn gekomen. Hieronder geef ik kort voor zowel de handgeschreven als de gedrukte documenten weer welke domeinen handgeschreven en welke domeinen gedrukte teksten bevatten en hoe de transcriptieprocedure voor beide soorten documenten in zijn werk ging.

Handgeschreven documenten

De domeinen waarvoor alleen handgeschreven teksten in het corpus zijn opgenomen, zijn Academie, Liefdadigheid, Privé en Religie. Het domein Economie bestaat voor een deel uit handgeschreven teksten en voor het domein Literatuur zijn twee van de gebruikte toneelstukken alleen in handschrift overgeleverd. Voor al deze documenten zijn dus de transcriptieconventies voor handgeschreven documenten gevolgd. De uitgebreide beschrijving van de transcriptieconventies voor handgeschreven documenten staat in bijlage 1.

Alle transcripties worden voorafgegaan door een header, waarin de belangrijkste metadata per transcriptie wordt gegeven: de bron van het originele document, het domein en het genre waartoe het document behoort, de datum waarop het originele document is geschreven en de plaats waar het is geschreven, het aantal pagina's van het originele document, het woordenaantal van de transcriptie (zonder de header), de initialen van de transcribent en de controlerende transcribent, en eventuele opmerkingen over de transcriptie, inclusief de naam van de auteur van het originele document, als die bekend is.

Van belang was dat de documenten diplomatisch getranscribeerd werden, dat wil zeggen dat de tekst van het originele document zo dicht mogelijk wordt gevolgd in de transcriptie. Dit betekent dat bijvoorbeeld spelling, interpunctie en het gebruik van hoofdletters en kleine letters precies zo worden weergegeven als in het originele document. Ook voor de pagina- en lijnafbrekingen werd het origineel gevolgd. Daarnaast is er gebruikgemaakt van verschillende tags in de transcriptie voor onduidelijke en onleesbare woorden (<ambig>woord</ambig> en <<illeg>>), doorhalingen (tekst en <<cancel>>) en woordafbrekingen aan het einde van een regel (<hyph>woo<->rd</hyph>). In sommige zestiende- en zeventiende-eeuwse teksten komen daarnaast afkortingen voor, bijvoorbeeld bij woorden met *ver* (*verschijnen*; *over*) en bij *ende*. Om het tekstbegrip te vergroten, zijn deze afkortingen opgelost, waarbij de tag <abb>tekst</abb> is gebruikt.

De transcripties zijn grotendeels gemaakt door student-assistent Jacqueline Maes en mijzelf. Ook Katharina Gunkler heeft in het kader van haar stage delen van de domeinen Academie en Religie getranscribeerd en Odette Pielage heeft voor haar BA-werkstuk transcripties gemaakt voor het Privédomein, die daarna ook in het *LOL*-corpus zijn opgenomen. Transcripties die al door verschillende studenten in het kader van een MA-werkgroep historische sociolinguïstiek waren gemaakt

en werden geselecteerd voor het *LOL*-corpus (voor het domein Privé), zijn aangepast aan de transcriptieconventies van het *LOL*-corpus. Alle eerste versies van de transcripties zijn daarna nog een keer nagekeken door mij of door Jacoline Maes om eventuele fouten in de transcriptie of inconsistenties in het gebruik van de transcriptieconventies recht te zetten. Vervolgens zijn de transcripties als definitief aangemerkt en opgenomen in het *LOL*-corpus.

Gedrukte documenten

Voor het domein Publieke Opinie zijn alleen gedrukte teksten in het *LOL*-corpus opgenomen; daarnaast zijn de meeste documenten uit het domein Literatuur en een deel van het domein Economie gedrukt. Al deze documenten zijn getranscribeerd volgens de transcriptieconventies voor gedrukte documenten. Deze zijn opgenomen in bijlage 2.

De transcripties worden, net als de transcripties voor handgeschreven documenten, voorafgegaan door een header met metadata. Ook voor de gedrukte documenten is er zo veel mogelijk diplomatisch getranscribeerd. Voor de transcripties van de kranten is echter niet de bladspiegel van het origineel gevolgd, waarbij de tekst in kolommen is verdeeld. In plaats daarvan zijn alle kolommen onder elkaar getranscribeerd. Wel zijn alle originele pagina- en regelafbrekingen aangehouden.

Vanwege het ontbreken van enkele verschijnselen die karakteristiek zijn voor handgeschreven teksten, zoals doorhalingen, waren er aanzienlijk minder tags nodig voor de gedrukte documenten. Wel kon het voorkomen dat een of meer woorden uit de gedrukte documenten niet goed te lezen waren, bijvoorbeeld door beschadigingen aan het papier. Daarom zijn de tags voor onduidelijke en onleesbare woorden (<ambig>woord</ambig> en <<illeg>>) ook gebruikt voor de transcriptieconventies van gedrukte documenten. Daarnaast was de tag voor woordaafbrekingen aan het einde van een regel (<hyph>woo<->rd</hyph>) van toepassing op gedrukte documenten.

Een groot deel van de transcripties voor Publieke Opinie was al gemaakt door studenten in het kader van verschillende MA-werkgroepen historische sociolinguïstiek; daarnaast heb ik zelf nog enkele aanvullende transcripties gemaakt. De transcripties zijn door student-assistent Jacoline Maes of mijzelf nagekeken op eventuele transcriptiefouten en aangepast aan de transcriptieconventies van het *LOL*-corpus. Hanna

Butter heeft in het kader van haar MA-scriptie de meeste achttiende- en negentiende-eeuwse toneelstukken voor het domein Literatuur opgespoord en getranscribeerd; Jacqueline Maes heeft de rest van de transcripties voor dit domein gemaakt. Al deze transcripties zijn ook nog een keer door mijzelf nagekeken voor ze aan het definitieve *LOL*-corpus werden toegevoegd.

4.3.3 *Corpus*

Het *LOL*-corpus zou idealiter een grootte hebben van 280.000 woorden: 5.000 woorden per periode van 50 jaar (waarvan er acht zijn in totaal) en per domein (waarvan er zeven zijn in totaal). Zoals in paragraaf 4.3.1 is beschreven, was het echter niet mogelijk om bij elk domein voor elke periode materiaal te vinden. Met name de eerste helft van de zestiende eeuw bleek problematisch, aangezien voor slechts twee van de zeven domeinen, Economie en Liefdadigheid, Leidse bronnen voor deze periode in de archieven konden worden gevonden. Voor de domeinen Academie, Publieke Opinie en Religie is hier een logische verklaring voor te geven: de universiteit bestond nog niet in de eerste helft van de zestiende eeuw, waardoor daar onmogelijk materiaal voor te vinden kan zijn; het krantengenre, gebruikt voor de Publieke Opinie, bestond pas vanaf de zeventiende eeuw; en voor het domein Religie is gebruik gemaakt van materiaal van protestantse kerken, die in Leiden pas vanaf 1572 bestonden. Alleen voor de domeinen Privé en Literatuur had er theoretisch gezien materiaal uit de periode 1500–1549 toegevoegd kunnen worden aan het corpus, maar dit soort materiaal is niet in de archieven aangetroffen.

Zoals gezegd bestaat het krantengenre pas vanaf de zeventiende eeuw, en de eerste Leidse krant werd pas vanaf de tweede helft van de zeventiende eeuw uitgegeven, waardoor het voor het domein Publieke Opinie niet alleen onmogelijk was om de periode 1500–1549 in te vullen, maar ook de perioden 1550–1599 en 1600–1649.

Behalve voor de periode 1500–1549 voor de domeinen Academie, Literatuur, Privé en Religie en de perioden 1500–1549, 1550–1599 en 1600–1649 voor het domein Publieke Opinie is voor elke combinatie van periode en domein geschikt materiaal gevonden voor het *LOL*-corpus. Voor bijna alle domeinen in alle perioden was er ook voldoende materiaal om het doel van 5.000 woorden te halen, behalve bij het Privédomein in de periode 1550–1599: hier kon slechts een aantal

van 4.449 woorden getranscribeerd worden, aangezien er niet meer materiaal aanwezig was.

Met de bovenstaande afwijkingen van het oorspronkelijke plan voor de opbouw van het *LOL*-corpus zijn de uiteindelijke woorden-aantallen per domein en periode als gegeven in tabel 4.1 op de volgende pagina. De domeinen in de tabel zijn gerangschikt op het totaal aantal cellen dat gevuld is, met de Publieke Opinie geheel links als het enige domein waar slechts vijf van de acht perioden zijn gevuld en de volledig gevulde perioden van de domeinen Economie en Liefdadigheid geheel rechts. Zoals te zien is, heeft het definitieve *LOL*-corpus een omvang van 251.417 woorden.

4.4 Corpusanalyses

Het *LOL*-corpus vormt het uitgangspunt voor de casestudies die in de volgende zes hoofdstukken worden besproken. Het gebruik van uit het Frans afkomstige suffixen is onderzocht (hoofdstuk 5 en 6), het gebruik van Franse leenwoorden (hoofdstuk 7, 8 en 9) en het gebruik van de betreffende voornaamwoorden *develke* en *betwelk* (hoofdstuk 10). Met de resultaten van de casestudies is het mogelijk om de invloed van het Frans op drie verschillende taalniveaus te onderzoeken: de morfologie, het lexicon en de (morfo)syntaxis. Voor elk van de onderzochte variabelen – leensuffixen, leenwoorden en betreffende voornaamwoorden – wordt, na een inleiding (paragraaf 1) en een bespreking van de bestaande literatuur over het onderwerp (paragraaf 2), de precieze methode besproken waarmee de data uit het corpus zijn gehaald (paragraaf 3). In paragraaf 4 volgen de resultaten van de analyses, waarbij aandacht wordt besteed aan de invloed van de variabelen periode en domein op het gebruik van de Franse elementen. Daarnaast zijn voor de suffixen en met name voor de leenwoorden enkele extra analyses uitgevoerd die bijdragen aan een zo precies mogelijk beeld van de daadwerkelijke invloed van het Frans op het Nederlands. Bij de leensuffixen is gekeken naar het verschil tussen algemene en domein-specifieke woordenschat en naar de mate van integratie van de suffixen in het Nederlands. Bij de leenwoorden is ook de algemene versus domein-specifieke woordenschat en de integratie van de leenwoorden onderzocht, maar daarnaast is er aandacht besteed aan de woordsoorten van de leenwoorden, het jaar van ontlening en de frequentie van individuele leenwoorden.

domein	Publieke Opinie	Privé	Academie	Literatuur	Religie	Liefdadigheid	Economie
genre	kranten- artikelen	brieven	notulen	toneelstukken	notulen	testamenten	ordonnanties verzoeken
1500–1549	n.v.t.	–	n.v.t.	–	n.v.t.	5.027	5.072
1550–1599	n.v.t.	4.449	5.046	5.116	5.305	5.229	5.118
1600–1649	n.v.t.	5.114	5.124	5.138	5.259	5.131	5.276
1650–1699	5.053	5.032	5.177	5.143	5.128	5.111	5.314
1700–1749	5.111	5.421	5.025	5.183	5.153	5.082	5.189
1750–1799	5.095	5.116	5.067	5.112	5.128	5.290	5.212
1800–1849	5.084	5.145	5.160	5.173	5.258	5.114	5.100
1850–1899	5.088	5.038	5.157	5.194	5.271	5.037	5.052
totaal	25.431	35.315	35.756	36.059	36.502	41.021	41.333
totaal aantal woorden: 251.417							

Tabel 4.1. Opbouw van het *Language of Leiden*-corpus met de aantallen woorden per domein en periode.

Omdat ik de leensuffixen- en leenwoordendataset zo uitgebreid heb geanalyseerd, zijn de resultaten die deze extra deelanalyses opleverden opgenomen in twee aparte hoofdstukken (hoofdstuk 6 voor de suffixen en hoofdstuk 7 en 8 voor de leenwoorden). Na elke casestudy volgt een korte conclusie waarin de resultaten worden gekoppeld aan de vraag wat de daadwerkelijke invloed van het Frans op het Nederlands is geweest. In de algemene conclusie in hoofdstuk 11 wordt een overkoepelend antwoord op deze vraag gegeven gebaseerd op alle casestudies samen.

Net als Krogull (2018) en Van de Voorde (2022) heb ik in mijn analyses de beschrijvende statistiek als uitgangspunt genomen, waarbij ik de kwantitatieve gegevens die mijn onderzoek opleveren in tabellen en grafieken weergeef. Beschrijvende statistiek heeft als voordeel ten opzichte van geavanceerde statistische methodes dat het minder tijdrovend is en dat er dus meer tijd is om verschillende variabelen te analyseren (Krogull 2018: 74). Door juist de voorkeur te geven aan beschrijvende statistiek was het mogelijk om drie verschillende taalniveaus in mijn onderzoek mee te nemen, waardoor een completer beeld kan worden gevormd van de daadwerkelijke invloed die het Frans op het Nederlands heeft gehad.