



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 2 Theoretisch kader en onderzoeksvragen

In dit proefschrift wordt de historische taalcontactsituatie tussen het Nederlands en het Frans onderzocht; meer specifiek bestudeer ik de invloed van het Frans op het Nederlands op het gebied van lexicon, morfologie en morfosyntaxis. Dit onderzoek is ingebed in het onderzoeksveld van de historische sociolinguïstiek, dat kort wordt geïntroduceerd in paragraaf 2.1.1. Om de invloed van het Frans op het Nederlands te onderzoeken, maak ik gebruik van het *LOL*-corpus, een historisch tekstcorpus dat ik speciaal voor dit doeleinde heb gebouwd. Ook ander historisch-sociolinguïstisch onderzoek maakt gebruik van corpora; enkele van de belangrijkste corpora die al bestaan bespreek ik in paragraaf 2.1.2. In tegenstelling tot al deze corpora is het corpus dat voor het huidige onderzoek wordt gebruikt opgebouwd op basis van de variabele ‘domein’; andere vergelijkbare corpora maken onderscheidingen op basis van de variabele ‘genre’. Het verschil tussen beide concepten en de overwegingen om voor domeinen te kiezen als basis voor de corpusopbouw en niet voor genres leg ik uit in paragraaf 2.2. Ook gaat deze paragraaf in op de definitie van ‘administratieve teksten’: het onderscheid tussen administratieve en niet-administratieve domeinen wordt in dit proefschrift veelvuldig gemaakt bij de diverse casestudies. In paragraaf 2.3 en 2.4 wordt vervolgens ingegaan op het concept ‘taalcontact’. Paragraaf 2.3 behandelt de taalcontacttheorie. In paragraaf 2.4 wordt ingezoomd op de specifieke Nederlands-Franse contactsituatie die in dit proefschrift centraal staat. Daarbij ga ik in op de bestaande kennis over deze contactsituatie die uit eerder onderzoek naar voren komt. Paragraaf 2.5 tot slot geeft de belangrijkste onderzoeksvragen die in dit proefschrift centraal worden gesteld en worden beantwoord.

2.1 De historische sociolinguïstiek

2.1.1 Het onderzoeksveld

Het onderzoek dat in dit proefschrift wordt gepresenteerd is historisch-sociolinguïstisch van aard. In Hernández-Campoy & Conde-Silvestre (2012: 1) wordt de historische sociolinguïstiek gedefinieerd als de reconstructie van de geschiedenis van een bepaalde taal in zijn sociaal-

culturele context. Als apart onderzoeksveld binnen de taalkunde bestaat de historische sociolinguïstiek nu zo'n veertig jaar: de studie *Socio-Historical Linguistics: Its Status and Methodology* van Suzanne Romaine (1982) wordt vaak als het beginpunt beschouwd van deze onderzoekstraditie. Romaine (2005: 1696) zelf en ook Auer e.a. (2015: 2) geven echter aan dat in de publicatie van Weinreich, Labov & Herzog (1968) weliswaar niet de term 'historische sociolinguïstiek' (of een aanverwante term) wordt gebruikt, maar dat deze publicatie desalniettemin gezien kan worden als het begin van het onderzoeksveld van de historische sociolinguïstiek, vanwege de nadruk die wordt gelegd op de historische en sociale benadering van taal. Maar ook daarvoor al was er oog voor de taalgeschiedenis en de relatie met sociale factoren, zoals in het werk van de Russische filosoof Bakhtin (Burke 2004: 4).

De historische sociolinguïstiek heeft een relatie met verschillende andere disciplines als de historische taalkunde, corpuslinguïstiek en sociale geschiedenis, maar leunt theoretisch en methodologisch het zwaarst op de (moderne) sociolinguïstiek (Hernández-Campoy & Conde-Silvestre 2012: 2; Nevalainen & Raumolin-Brunberg 2012: 23–27). De historische sociolinguïstiek maakt gebruik van de theoretische principes van de sociolinguïstiek, die het hedendaagse taalgebruik in zijn sociale context bestudeert, en past die toe op de taal van het verleden (Hernández-Campoy & Conde-Silvestre 2012: 1). Dit gebeurt op basis van het *uniformitarian principle*, dat zegt dat de taalveranderingsmechanismen die in het heden geobserveerd kunnen worden ook toegepast kunnen worden op de taalveranderingsmechanismen in het verleden, en andersom (Labov 1972: 275; Hernández-Campoy & Conde-Silvestre 2012: 5; Nevalainen & Raumolin-Brunberg 2012: 24–25; Auer e.a. 2015: 4; zie voor een uitgebreide analyse van de betekenis van het begrip *uniformitarian principle* in de taalwetenschap Walkden 2019). Er moet hierbij echter wel rekening worden gehouden met het feit dat bepaalde concepten als 'sociale klasse' en 'gender' en bijvoorbeeld genreconventies zoals die op de moderne, westerse maatschappij van toepassing zijn niet zonder meer op hoeven te gaan voor de samenlevingen van het verleden. De historische sociolinguïstiek kan deze concepten daardoor niet zonder meer van de moderne sociolinguïstiek overnemen, maar moet deze hanteren binnen de sociale context van het verleden, die daarvoor gereconstrueerd moet worden (Auer e.a. 2015: 5; Nevalainen & Raumolin-Brunberg 2012: 29). Daarnaast kan de historische sociolinguïstiek geen gebruik maken van hetzelfde materiaal als de moderne sociolinguïstiek. De eerste kan zich, bij gebrek aan

opnames van gesproken taal van voor de twintigste eeuw, alleen op geschreven bronnen baseren om de taal van het verleden te bestuderen. Daarbij is de onderzoeker afhankelijk van het overgeleverde tekstmateriaal, dat veelal eenzijdig is en vooral formeel taalgebruik van mannen uit de hogere klassen omvat. Daardoor blijven veel groepen van de samenleving onderbelicht (Auer e.a. 2015: 5–6). De ‘*language history from below*’-benadering binnen de historische sociolinguïstiek wil hier verandering in brengen door gebruik te maken van materiaal waarin ook het taalgebruik van de lagere klassen is gerepresenteerd, zoals privé-brieven (Elsaß 2012: 160).

Er zijn veel verschillende onderwerpen die binnen de historische sociolinguïstiek worden onderzocht, zoals de invloed van sociale factoren als sociale klasse, gender, leeftijd en sociale netwerken op taalveranderingen in het verleden en op de verspreiding van de taalveranderingen in ruimte en tijd, maar ook bijvoorbeeld de invloed van standaardisatie op taalverandering en taalvariatie (Hernández-Campoy & Conde-Silvestre 2012: 3). Ook historisch taalcontact en meertaligheid is een tak van de historische sociolinguïstiek waarbinnen veel onderzoek wordt gedaan. Met name naar de middeleeuwse en vroegmoderne taalcontactsituatie in het Engelse taalgebied, onder andere het taalcontact met het Latijn en het Frans, is al veel onderzoek gedaan (Schendl 2012: 520). Bij deze onderzoekstak van de historische sociolinguïstiek sluit het onderzoek in dit proefschrift aan.

Daarnaast sluit dit proefschrift aan bij de historisch-sociolinguïstische onderzoekstraditie voor de Nederlandse taal. Historisch-sociolinguïstisch onderzoek voor het Nederlands wordt met name uitgevoerd aan de Universiteit Leiden en de Vrije Universiteit Brussel. In het project *Brieven als Buit* is aan de Universiteit Leiden onderzoek gedaan naar de rol van de factoren sociale klasse, gender, leeftijd en regio in het taalgebruik in zeventiende- en achttiende-eeuwse privé-brieven uit de Noordelijke Nederlanden (zie Rutten & Van der Wal 2014). Het project *Going Dutch* aan dezelfde universiteit bracht de implementatie van taalbeleid van het Nederlands in het onderwijs en het effect daarvan op het taalgebruik in de Noordelijke Nederlanden aan het begin van de negentiende eeuw in kaart (zie Rutten 2019). Het onderzoek aan de Vrije Universiteit Brussel is primair gericht op de historische sociolinguïstiek van de Zuidelijke Nederlanden/Vlaanderen. Voorbeelden van projecten zijn de standaardisatie van het Nederlands in Vlaanderen in de negentiende eeuw en de invloed van sociale factoren op het taalgebruik in achttiende- en negentiende-eeuwse egodocumenten uit Vlaanderen (zie

Vandenbussche 2007; Vosters & Vandenbussche 2012; Vosters e.a. 2014). Naar contact met andere talen is echter binnen de Nederlandse historische sociolinguïstiek nog relatief weinig onderzoek gedaan.¹ Hier wil dit proefschrift verandering in brengen.²

2.1.2 *Historische corpora*

De dominante onderzoeksmethode die in de historische sociolinguïstiek wordt gebruikt is corpusonderzoek. Digitale, historische tekstcorpora spelen een belangrijke rol in met name kwantitatief historisch-sociolinguïstisch onderzoek (Nevalainen & Raumolin-Brunberg 2012: 26; Auer e.a. 2015: 6). Ook voor het onderzoek dat in dit proefschrift is beschreven, is gebruik gemaakt van een corpus, waarin de variabelen ‘domein’ en ‘periode’ centraal staan. Dit is hiermee het eerste historisch-sociolinguïstische ‘multi-domein’-corpus. Andere corpora hebben in plaats van ‘domein’ ‘genre’ als basis voor de corpusopbouw, of hebben slechts één genre/domein in het corpus opgenomen. In paragraaf 2.2 hierna ga ik in op het verschil tussen de variabelen domein en genre en de overwegingen om in het huidige onderzoek domein als variabele te nemen in plaats van genre. In de rest van de huidige paragraaf beschrijf ik de belangrijkste historische corpora die gedurende de laatste veertig jaar zijn samengesteld voor historisch-sociolinguïstisch onderzoek. Ik beperk mij hier tot een bespreking van Engelse, Duitse en Nederlandse corpora.

Uit de begintijd van de historische sociolinguïstiek stamt het multi-genrecorpus *Helsinki corpus of English texts*, dat in de jaren tachtig is ontwikkeld aan de universiteit van Helsinki.³ Het diachrone deel van het corpus⁴ beslaat een tijdspanne van zo’n 1.000 jaar, van circa 750 tot

¹ Vanhecke (2007) en Kessels-Van der Heijde (2002) behandelen bijvoorbeeld wel taalcontact tussen het Nederlands en andere talen, met name het Frans, maar dan niet vanuit een sociolinguïstische maar vanuit een taalsociologische invalshoek (taalkeuze).

² Ook aan de Vrije Universiteit Brussel loopt momenteel een onderzoeksproject, uitgevoerd door promovenda Charlotte Verheyden, naar taalcontact tussen het Nederlands en het Frans, maar dan in de Zuidelijke Nederlanden. De titel van het project is: *Hadde ik maer fransch kunnen spreken. Een historisch-sociolinguïstisch onderzoek naar de invloed van het Frans op het Zuidelijke Laatinieuwnederlands*.

³ Het corpus is online beschikbaar via <https://helsinki.corpus.arts.gla.ac.uk/> (geraadpleegd op 18 augustus 2022).

⁴ Er is ook een dialectisch deel, dat bestaat uit hedendaagse Britse plattelandsdialecten (Rissanen 1994: 73).

1700. Deze tijdspanne is verdeeld in perioden met een lengte van een eeuw voor de vroegste teksten en van 70 of 80 jaar voor de latere teksten (Kytö & Rissanen 1992: 7–8). De teksttypes die zijn opgenomen in het corpus zijn zeer divers: er zitten onder andere wetteksten, wetenschappelijke handboeken, filosofische teksten, preken, historische teksten, reisverslagen, biografieën, romans, toneelstukken, brieven en Bijbelteksten in het corpus (Kytö & Rissanen 1992: 15). Ook is een deel van het corpus gecodeerd voor dialect (Kytö & Rissanen 1992: 17–18). Het totaal aantal woorden is circa 1,5 miljoen (Rissanen 1994: 73).

Het *ARCHER*-corpus (*A Representative Corpus of Historical English Registers*) wordt sinds begin jaren negentig ontwikkeld aan verschillende universiteiten. Het is een multi-genrecorpus met Brits-Engelse teksten uit de periode 1600–1999 en Amerikaans-Engelse teksten uit de periode 1750–1999 en het bevat in de laatste versie ongeveer 3,3 miljoen woorden. De teksten zijn opgedeeld in perioden van elk vijftig jaar, een tijdspanne die ik voor mijn corpus ook hanteer (zie hoofdstuk 4). Het corpus bevat verschillende registers/genres als dagboeken, privébrieven, nieuwsreportages, juridische adviezen, medische teksten, wetenschappelijke teksten, preken en dialogen in toneelstukken en fictie (*ARCHER* 2013).⁵ Op deze manier zijn zowel formele als informele schrijfstijlen (bijv. juridische adviezen en preken als formeel en dagboeken en brieven als informeel) en zowel teksten gebaseerd op de gesproken taal als de geschreven taal (bijv. preken en dialogen als gesproken taal en wetenschappelijke teksten als geschreven taal) in het corpus gerepresenteerd (Yáñez-Bouza 2011: 207).

Een single-genrecorpus voor het Engels dat ook in de jaren negentig is ontwikkeld is *The Helsinki Corpus of Early English Correspondence* (*CEEC*). Het corpus bestaat uit privébrieven uit de periode 1410–1681 en bevat zo'n 2,7 miljoen woorden (Nevalainen & Raumolin-Brunberg 2016: 29–30). Later is dit corpus nog uitgebreid, onder andere met brieven uit de periode 1681–1800 (Kaislaniemi 2018: 48).⁶ Het *CEEC* was het eerste corpus waarin uitgebreide informatie over de sociale factoren van de briefschrijvers als sociale klasse, gender en leeftijd werd opgenomen (Auer e.a. 2015: 6).

⁵ Toegang tot het corpus is aan te vragen via <https://www.projects.alc.manchester.ac.uk/archer/archer-versions/> (geraadpleegd op 27 juni 2022).

⁶ Zie <https://www2.helsinki.fi/en/researchgroups/varieng/corpus-of-early-english-correspondence> (geraadpleegd op 27 juni 2022) voor de link naar het corpus.

Voor het Duits is er het *GerManC*, een historisch multi-genrecorpus dat is ontwikkeld aan de Universiteit van Manchester. Het corpus bestaat uit geschreven Duitse teksten uit de periode 1650 tot 1800 opgedeeld in drie perioden van elk 50 jaar. Het corpus representeert verschillende tekstgenres, waaronder toneelstukken, juridische teksten, brieven, kranten, wetenschappelijke teksten en preken. Ook regio is als variabele in het corpus opgenomen. Het totaal aantal woorden bedraagt 800.000 (Durrell e.a. 2012).⁷ Eveneens voor het Duitse taalgebied bestaat een corpus negentiende-eeuwse emigrantebrieven van in totaal 375.000 woorden, met name geschreven door mannen en vrouwen uit de lagere klassen die zich vanuit verschillende Duitse regio's in de Verenigde Staten hadden gevestigd (Elspaß 2007).

Voor het Nederlands bestaat het eerder genoemde *Brieven als Buit*-corpus met 424.000 woorden uit de late zeventiende en late achttiende eeuw, geschreven door mannen en vrouwen uit verschillende sociale klassen, van verschillende leeftijden en uit verschillende regio's.⁸ Recenter zijn er twee diachrone multi-genrecorpora voor de Nederlandse taal opgesteld. Aan de Universiteit Leiden is tussen 2014 en 2018 het *Going Dutch corpus* opgebouwd met teksten uit twee perioden (1770–1790 en 1820–1840). Elke periode is gerepresenteerd in het corpus met teksten uit verschillende regio's en drie genres (privébrieven, reisverslagen/dagboeken en kranten). Ook is bij de privébrieven en de dagboeken gender als variabele meegenomen in het corpus. Het totale corpus bestaat uit zo'n 420.000 woorden (zie Krogull 2018).

Het tweede Nederlandse multi-genrecorpus, het *HCD* (*Historical Corpus of Dutch*), is tussen 2017 en 2021 aan de Vrije Universiteit Brussel en de Universiteit Leiden samengesteld. Dit corpus bevat taalmateriaal uit de zestiende tot en met de negentiende eeuw, waarmee dit corpus in tijdspanne overeenkomt met het corpus dat voor dit proefschrift wordt gebruikt. Voor het *HCD* is voor elke eeuw een periode van veertig jaar rond het midden van die eeuw genomen (dus 1530–1570, 1630–1670, 1730–1770 en 1830–1870). Naast verschillende regio's werden er drie genres geselecteerd, te weten administratieve teksten, egodocumenten en pamfletten. Het totaal aantal woorden van het *HCD* is 460.000 (zie Van de Voorde 2022; Van de Voorde e.a. te verschijnen).

⁷ Het corpus is online beschikbaar via https://www.laudatio-repository.org/browse/corpus/9SX7DnMB7CArCQ9CBB_J/corpora (geraadpleegd op 18 augustus 2022).

⁸ Het corpus is online beschikbaar via <https://brievenalsbuit.ivdnt.org/> (geraadpleegd op 7 juli 2022).

In hoofdstuk 4 bespreek ik de opbouw van het *Language of Leiden*-corpus, dat ik voor het huidige onderzoek heb samengesteld. Dit corpus vertoont overeenkomsten met enkele van bovengenoemde corpora op het gebied van periodisering en selectie van genres voor de domeinen, maar in tegenstelling tot bovenstaande corpora is het corpus samengesteld op basis van de variabele domein en ook is het corpus opgebouwd uit teksten uit één locatie: Leiden. In hoofdstuk 4 (paragraaf 4.1) ga ik in op de vraag waarom ervoor is gekozen om het onderzoek op Leiden te richten; in de hierna volgende paragraaf licht ik de keuze voor de variabele domein in plaats van genre toe.

2.2 Domeinen en genres

Zoals gezegd wordt in dit proefschrift gewerkt met het begrip domein: het is een van de twee variabelen op basis waarvan het *LOL*-corpus is samengesteld. ‘Domeinen’ kunnen kort worden aangeduid als ‘verschillende gebieden van menselijke activiteit’ (Ehala 2019: 536). Fishman (1972b: 452) definieert een domein als een ‘cluster of social situations typically constrained by a common set of behavioral rules’. Een domein is een sociale setting, situatie, werkterrein of institutie zoals familie, school, kerk of werk, waarin één of meer talen worden gebruikt (Ehala 2019: 536). Het concept ‘domein’ wordt typisch gebruikt om taalcontact- of meertaligheidssituaties te beschrijven, waarbij de achterliggende gedachte is dat de keuze van een zekere taal of taalvariëteit door meertalige sprekers in een meertalige maatschappij afhankelijk is van het domein waarin de communicatie plaatsvindt (Fishman 1972b: 435–436; Uribe-Villegas 1977: 18; Haberland 2019: 397). Het gebruik ervan gaat terug tot de Duitse wetenschapper Schmidt-Rohr (1933). Hij onderscheidde negen domeinen waarin hij de verhouding onderzocht tussen het gebruik van Duitse dialecten en de nationale standaardtalen in landen waar Duits een minderheidstaal was (zoals Italië en Hongarije). Deze negen domeinen waren de familie, de speelplaats, de school, de kerk, literatuur, kranten, het leger, de rechtbanken en het openbaar bestuur. Fishman (1964, 1965, 1972b) werkte het terrein van de domeinanalyse verder uit en introduceerde dit concept in de sociolinguïstiek om meertalige taalgemeenschappen te analyseren. Dit soort onderzoeken zijn echter meer taalsociologisch van aard en richten zich met name op taalkeuze in meertalige contexten.

Greenfield (1968) en Fishman (1968), beide geciteerd in Fishman (1972a: 22–26), onderscheidden vijf hoofddomeinen: familie, vrienden, religie, school en werk, maar Fishman (o.a. 1965: 74) merkte ook op dat het afhankelijk is van de cultuur welke domeinen (daarnaast) relevant zijn. Domeinen zijn een theoretische notie en zijn niet direct waarneembaar in de maatschappij, maar er bestaat wel een zekere verwachting bij gespreksdeelnemers met betrekking tot de rolverdeling, de setting en de onderwerpen waarover gecommuniceerd wordt in een zeker domein, waardoor de domeinen geïdentificeerd en van elkaar onderscheiden kunnen worden. Het doel van een domeinanalyse is om patronen van taalkeuze te ontdekken in bepaalde domeinen (Haberland 2019: 399).

Een historisch-sociolinguïstische studie die gebruikmaakt van het domeinconcept, is Kessels-Van der Heijde (2002; 2015). De studie onderzoekt de taalsituatie in Maastricht in de negentiende eeuw, waar zowel Nederlands als Maastrichts en Frans werden gebruikt. Kessels-Van der Heijde (2002) onderscheidt in haar onderzoek zeven domeinen: administratie, onderwijs, geschreven media, commerciële instellingen, familie, toneel en verenigingen en laat zien dat de verhoudingen in het gebruik van de drie talen per domein verschillen. Zo domineerde het Frans de Maastrichtse familiearchieven tot het einde van de negentiende eeuw, terwijl er weinig Nederlands materiaal en vrijwel geen Maastrichts werd gevonden, maar in het domein van de verenigingen werd er gedurende de negentiende eeuw juist steeds meer Maastrichts gebruikt, ten koste van het Frans (Kessels-Van der Heijde 2002: 289–292).

In het corpus dat in dit proefschrift wordt gebruikt, wordt een domein meestal gerepresenteerd door één genre. Genres en registers hebben onder andere situationele kenmerken die ze van andere genres en registers onderscheiden. Onder situationele kenmerken vallen de gespreksdeelnemers (wie zijn ze, met hoe veel zijn ze, etc.) en de onderlinge relaties tussen de gespreksdeelnemers, de manier waarop de communicatie plaatsvindt (geschreven of gesproken, door welk medium), de setting van de communicatie in termen van tijd en plaats, de communicatieve doelen van de communicatie en het onderwerp (Biber & Conrad 2009: 40). Wanneer een tekst op deze kenmerken verschilt van een andere tekst, dan behoren de teksten tot verschillende genres of registers. Een genre en een register onderscheiden zich dan weer van elkaar doordat bij een genreperspectief de nadruk ligt op specifieke talige kenmerken die gebruikt worden om de tekst te structureren (zoals een aanhef en een afsluiting bij een brief) terwijl bij het registerperspectief de nadruk ligt op taalkenmerken die veel voorkomen in dat specifieke

register en ook een communicatieve functie vervullen (zoals het gebruik van bepaalde voornaamwoorden of werkwoordstijden) (Biber & Conrad 2009: 16).

Hoewel teksten op basis van de situationele kenmerken tot verschillende genres kunnen behoren, kunnen deze echter nog wel onder hetzelfde sociale domein vallen: één domein kan geassocieerd worden met meer dan één genre. Zo zijn besluiten, reglementen, wetten en ordonnanties voorbeelden van genres die in het bestuurlijke domein voorkomen, terwijl gebeden, preken en zegeningen tot het religieuze domein kunnen worden gerekend (Biber & Conrad 2009: 34). Omgekeerd kan één genre ook bij verschillende domeinen horen. In het *LOL*-corpus zijn bijvoorbeeld bij twee domeinen – Academie en Religie – notulen van vergaderingen gebruikt. Een onderscheiding op basis van domeinen maakt het dus mogelijk om het taalgebruik van hetzelfde genre in verschillende situaties te bestuderen. Omdat dit onderzoek gericht is op een taalcontactsituatie – met het Nederlands en het Frans als contacttalen – en domeinonderscheidingen juist bij onderzoek naar dit soort settings worden gebruikt, ligt de keuze voor een analyse op basis van de variabele domein in plaats van genre ook voor de hand. Weliswaar is er tot nu toe vooral gebruikgemaakt van domeinonderscheidingen in onderzoek naar taalkeuze, maar ook voor sociolinguïstisch onderzoek kan deze variabele relevant zijn; dit is iets wat in dit proefschrift wordt getoetst. Tot slot was het doel om een corpus te maken dat representatief zou zijn voor de historische situatie van de stad Leiden. Door het corpus in te delen op basis van domein kon beter recht worden gedaan aan alle aspecten van de historische Leidse samenleving dan bij een genreonderscheiding mogelijk is. Om deze redenen is het *LOL*-corpus op basis van domein en niet op basis van genre opgebouwd.

Administratief versus niet-administratief

In dit proefschrift zal herhaaldelijk een onderscheid worden gemaakt tussen administratieve teksten en niet-administratieve teksten. Onder administratieve teksten vallen de teksten uit de domeinen Academie (notulen), Economie (besluiten en verzoeken), Liefdadigheid (testamenten) en Religie (notulen). Onder de niet-administratieve teksten vallen de domeinen Literatuur (toneelstukken), Privé (brieven) en Publieke Opinie (kranten).

Een definitie van (lokale) administratieve teksten wordt gegeven door Stenroos (2020: 43), die deze teksten schaaft onder de zogeheten *documentary texts*. Dit zijn teksten die een pragmatische functie hebben in plaats van een esthetische, wetenschappelijke of didactische functie: ze willen bijvoorbeeld iets gedaan krijgen (van de lezer), een beslissing vastleggen, etc. Daarnaast zijn *documentary texts* verbonden met een specifieke situatie op een specifiek moment, waarbij specifieke mensen zijn betrokken. Daardoor vallen bijvoorbeeld wetteksten niet onder de *documentary texts*, want deze zijn niet aan een specifieke tijd en plaats gebonden en hebben een algemene en geen specifieke toepassing (Stenroos 2020: 43).

Naast administratieve teksten kunnen volgens Stenroos (2020: 43) ook veel persoonlijke brieven volgens de bovenstaande definitie bij de *documentary texts* worden gerekend. Deze kunnen echter vanwege hun persoonlijke aard van de administratieve teksten worden onderscheiden. Voorbeelden van administratieve teksten die Stenroos (2020) geeft zijn teksten geproduceerd door onder andere de stedelijke overheid, gilden, de academie, kerken, de rechtbank en lokale hoven. Teksttypen die hierbij horen zijn bijvoorbeeld ordonnanties, verslagen, verkopen, testamenten en inventarissen. Deze teksttypen komen overeen met de teksten uit de domeinen Academie, Economie, Liefdadigheid en Religie in het *LOL*-corpus.

Een ander belangrijk onderscheid tussen de groepen ‘administratief’ en ‘niet-administratief’ is de mate van formaliteit van de teksten, waarbij de administratieve teksten als formeler kunnen worden gezien dan de niet-administratieve teksten. In het *ARCHER*-corpus is een onderscheid gemaakt tussen formele en informele teksten waarbij brieven en toneeldialogen als ‘informeel’ werden bestempeld, terwijl juridische teksten als ‘formeel’ werden gezien en nieuwsreportages tussen formeel en informeel in zaten (Yáñez-Bouza 2011: 207). De genres die in het *LOL*-corpus voor de domeinen zijn gebruikt kunnen op een vergelijkbare manier worden ingedeeld: privébrieven en toneelstukken voor de domeinen Privé respectievelijk Literatuur kunnen dan onder de informele teksten worden geschaard en de administratieve teksten uit de domeinen Academie, Economie, Liefdadigheid en Religie onder de formele teksten, terwijl de krantenartikelen uit Publieke Opinie een middenpositie innemen.

2.3 Taalcontact

In dit proefschrift worden taalveranderingen onderzocht die het gevolg zijn van taalcontact, een situatie waarbij mensen die verschillende talen spreken met elkaar communiceren, wat uiteindelijk leidt tot taalverandering in de betrokken talen. Myers-Scotton (2002: 30) beschrijft de verschillende stadia van taalcontact. Eerst komen sprekers van verschillende talen met elkaar in contact; om te kunnen communiceren, moeten zij een andere taal leren. Een mogelijke uitkomst hiervan is tweetaligheid: de spreker behoudt zijn vaardigheid in zijn eerste taal, maar krijgt ook een of meer tweede talen in zijn talige repertoire, die hij in meer of mindere mate beheerst. Een andere mogelijkheid is taalshift, waarbij de tweede taal de dominante taal wordt en de vaardigheid in de eerste taal langzaam maar zeker vervaagt.

Bepaalde omstandigheden zorgen voor de ontwikkeling van tweetaligheid, waaronder militaire verovering en/of kolonisatie, de nabijheid van een taalgrens, migratie, onderwijs en de verspreiding van internationale talen (*lingua franca*) (Myers-Scotton 2002: 31–32; Thomason 2001: 17–21). Daarnaast moeten de sprekers gemotiveerd zijn om een tweede taal te leren. Vaak komt die motivatie voort uit sociaaleconomische overwegingen, bijvoorbeeld omdat het voor het goed uitvoeren van een bepaalde baan of voor de baankansen nodig is om een andere taal te leren (Myers-Scotton 2002: 36–38; Thomason 2001: 49). Een andere reden om een nieuwe taal te leren heeft te maken met sociale identiteit: individuen willen zichzelf bijvoorbeeld met een bepaalde levensstijl of groep identificeren, waarbij het spreken van een bepaalde taal als bijvoorbeeld modern of beschaafd wordt gezien of als noodzakelijk om als volwaardig lid van een groep mee te tellen (Myers-Scotton 2002: 38–39; Thomason 2001: 48). Maar ook bij een huwelijk tussen sprekers van verschillende talen moet ten minste een van de twee een andere taal leren om het mogelijk te maken om met elkaar te communiceren (Thomason 2001: 48–49).

Een effect van tweetaligheid is dat het zorgt voor taalverandering in de vorm van substraatwerking of ontlening.⁹ Bij substraatwerking brengen sprekers elementen als syntactische structuren of de uitspraak van bepaalde klanken over van hun moedertaal op hun tweede taal (Van Bree 2000: 23). Ontlening houdt de overname in van morfemen en/of

⁹ Andere consequenties van meertaligheid zijn bijvoorbeeld code-switching, taalkeuze en taalshift, maar deze effecten van taalcontact blijven in dit proefschrift buiten beschouwing.

structuren van de tweede taal in de moedertaal (Myers-Scotton 2002: 41; Thomason 2001: 62; Matras 2009: 146; Van Bree 2000: 23). Aangezien in dit proefschrift taalverandering door ontlening centraal staat, beperk ik mij in de rest van deze paragraaf tot de bespreking van de theorie over ontleningen.

Een reden die in de taalcontactliteratuur wordt genoemd waarom taalgebruikers elementen uit een andere taal overnemen is om ‘gaten’ in de ontlenende taal op te vullen, waarbij bepaalde manieren om iets uit te drukken die wel bestaan in de taal waaraan ontleend wordt maar niet bestaan in de ontlenende taal in deze laatste worden geïntroduceerd. Ook het prestige van de taal waaruit ontleend wordt is een belangrijke motivatie voor ontlening, waarbij sprekers van de ontlenende taal elementen overnemen van taalgemeenschappen die sociaal dominanter zijn om op die manier de sociale status van de ontlenende partij te verhogen (Matras 2009: 149–150; Matras & Adamou 2021: 243). Een derde motivatie die vooral opgaat voor de ontlening van structuur-elementen en niet voor lexicale elementen is cognitief van aard: om de taalverwerking van een tweetalige spreker te versimpelen en de cognitieve last om een keuze te maken uit de verschillende grammaticale structuren die beschikbaar zijn in het repertoire van de tweetalige spreker te verlichten, wordt het aantal keuzes gereduceerd, met als gevolg dat de structuren van beide talen convergeren (Matras 2009: 151).

Volgens Thomason (2001: 63) zou vanuit een puur taalkundig oogpunt in principe alles uit een andere taal geleend moeten kunnen worden, maar volgens de taalcontactliteratuur zitten er wel beperkingen op de mogelijkheid voor taalverandering als gevolg van taalcontact. Deze beperkingen worden vaak weergegeven in zogeheten ontleningshiërarchieën. Deze hiërarchieën kunnen volgens Seifart (2019: 15) drie dingen doen: (1) voorspellen welke ontleningen mogelijk zijn in een individuele taal, bijvoorbeeld dat een taal pas een bepaald morfologisch element kan lenen als er eerst een ander morfologisch element is geleend; (2) diachrone tendensen verklaren, waarbij het ene element eerder in de tijd wordt geleend dan het andere element; (3) beschrijven wat de relatieve frequenties zijn van de geleende elementen ten opzichte van elkaar in typologische samples.

Er zijn verschillende soorten ontleningshiërarchieën, die elk rekening houden met verschillende kenmerken van de geleende elementen. Sommige hiërarchieën richten zich op formele kenmerken, waarbij bijvoorbeeld wordt gesteld dat een vrij morfeem makkelijker te ontlenen is dan een gebonden morfeem, een lexicaal item makkelijker

dan een niet-lexicale, grammaticale eigenschap en een zelfstandig naamwoord makkelijker dan elke andere woordsoort (Matras 2009: 155; Seifart 2019: 15). Andere hiërarchieën richten zich op semantische of pragmatische kenmerken, waarbij het idee is dat bijvoorbeeld morfemen met een concrete, lexicale betekenis eerder worden ontleend dan morfemen met een abstracte, grammaticale betekenis. Deze twee soorten hiërarchieën correleren vaak met elkaar in de zin dat bijvoorbeeld gebonden morfemen (een formeel kenmerk) vaak een abstracte, grammaticale betekenis hebben (een semantisch kenmerk) (Matras 2009: 155; Seifart 2019: 15–16). Een belangrijk onderscheid hierbij is ook het verschil tussen flexionele en lexicale morfologie: flexionele morfemen zijn sterker gebonden morfemen dan bijvoorbeeld de leensuffixen die in dit proefschrift worden behandeld, waardoor deze laatste groep morfemen ook sneller ontleend kan worden (Seifart 2019: 17).

Leenwoorden zijn vrije morfemen met vaak een concrete betekenis, wat verklaart waarom dit soort elementen uit andere talen wordt overgenomen. Woorden zouden ook de eerste elementen zijn die uit andere talen worden overgenomen en daarna pas grammatica (Thomason & Kaufman 1992: 37; Thomason 2001: 69). Ook binnen de categorie van leenwoorden zijn echter hiërarchieën te maken op basis van de relatieve mate van ontlening van de verschillende woordsoorten. Zo zijn er studies die laten zien dat zelfstandige naamwoorden eerder worden ontleend dan werkwoorden, werkwoorden eerder dan voorzetsels, voorzetsels eerder dan voornaamwoorden, et cetera (Matras 2009: 157; Seifart 2019: 16).

Thomason (2001: 70–71) geeft een ontleningsschaal gebaseerd op de intensiteit van het taalcontact. Alleen als er voor een langere periode een sterke culturele druk is van sprekers van de taal die de woorden uitleent (de brontaal) op de sprekers van de taal die de woorden leent (de ontvangende taal) en als er veel tweetalige sprekers zijn van de brontaal en de ontvangende taal, dan kunnen naast woorden ook andere elementen worden ontleend, zoals fonologische, fonetische of syntactische elementen (Thomason & Kaufman 1992: 37; Thomason 2001: 66; Seifart 2019: 18). Daarnaast speelt de mate van integratie van de te lenen elementen een rol in deze ontleningsschaal, wat overeenkomt met de hiërarchieën gericht op formele kenmerken die hierboven zijn genoemd. Dit betekent dat een zeer geïntegreerde, gesloten structuur als een verbuigingsuffix vanwege zijn integratie moeilijker te ontleen is dan bijvoorbeeld woordvolgorde (Thomason 2001: 69).

De ontleningsschaal van Thomason (2001: 70–71) bestaat uit vier niveaus. Op het eerste niveau is er informeel taalcontact, waarbij de sprekers van de ontlenende taal – de taalgebruikers dus die elementen ontleen uit de brontaal – de brontaal niet vloeiend spreken en/of waarbij er weinig tweetalige sprekers zijn onder de sprekers van de ontlenende taal. Op dit niveau worden alleen inhoudswoorden ontleend, vaak zelfstandige naamwoorden, maar ook werkwoorden, bijvoeglijke naamwoorden en bijwoorden. Op het tweede niveau is het taalcontact iets intenser, zijn de taalgebruikers de brontaal redelijk machtig maar vormen deze tweetalige sprekers een minderheid onder de sprekers van de ontlenende taal. In deze fase worden naast inhoudswoorden ook functiewoorden ontleend, bijvoorbeeld voegwoorden. Op het gebied van de structurelementen zijn er alleen kleine ontleningen en worden er bijvoorbeeld nieuwe fonemen en bijbehorende spraakklanken in de ontlenende taal geïntroduceerd. Op het derde niveau is er nog intenser taalcontact, met meer tweetalige sprekers. Er worden meer functiewoorden geleend en ook woorden uit de basiswoordenschat – de woorden die iedere taal heeft, bijvoorbeeld aanduidingen voor lage getallen – en derivatieve affixen kunnen worden ontleend. Structurelementen worden ook meer geleend, hoewel deze niet leiden tot grote typologische veranderingen in de ontvangende taal. Een voorbeeld is het toe- of afnemen van het gebruik van deelwoordconstructies in plaats van het gebruik van bijzinnen ingeleid door een voegwoord. Tot slot is er op het vierde niveau sprake van intens taalcontact, waarbij er heel veel tweetalige sprekers zijn. In dit stadium wordt er heel veel lexicaal materiaal geleend en kunnen alle structurele elementen uit de brontaal worden overgenomen, ook als dat betekent dat de ontlenende taal typologisch verandert. Er kunnen bijvoorbeeld hele fonologische categorieën bijkomen of verdwijnen in inheemse woorden en er kunnen permanente woordvolgordeveranderingen optreden.

2.4 Taalcontact met het Frans

De specifieke historische taalcontactsituatie die in dit proefschrift wordt onderzocht, is die tussen het Frans en het Nederlands. De contacten tussen het Frans en het Nederlands zijn al eeuwenoud. De beide taalgebieden grenzen aan elkaar; vanaf de middeleeuwen was deze taalgrens min of meer stabiel en liep grof gezegd van west naar oost door het huidige Noord-Frankrijk (tussen Calais en Duinkerke), Vlaanderen

en Brabant tot aan Aken (Armstrong 1965: 387–389; Van Durme 2002; Willemyns 2015: 40–41; Van de Haar 2019: 40–41). Deze nabijheid van het Franse taalgebied zorgde voor intens taalcontact tussen het Nederlands en het Frans (Peersman, Rutten & Vosters 2015: 1). Dit taalcontact kwam, afhankelijk van de periode, in verschillende domeinen tot stand.

In de middeleeuwen was er in een beperkt aantal domeinen contact met het Frans: voornamelijk op het gebied van bestuur, handel en diplomatie. Van 1299 tot 1354 was onder andere het graafschap Holland deel van het Huis van Avesnes, waardoor het Frans naast het Nederlands een belangrijke hoftaal werd in Holland. In de vijftiende eeuw werd Holland – samen met grote delen van de rest van het Nederlandstalige gebied – onderdeel van het Bourgondische en daarna van het Habsburgse rijk, waar het Frans ook een belangrijke administratieve taal was (Van de Haar 2019: 42–43; Van Strien-Chardonneau 2014: 145). Bovendien was het Frans vanaf de middeleeuwen in gebruik als de taal van diplomatie en internationale handel in de Nederlanden vanwege de grote hoeveelheid handel die met Frankrijk werd gedreven (Frijhoff 2015: 116).

Ten tijde van de Republiek der Zeven Verenigde Nederlanden (vanaf eind zestiende eeuw) bleef het Frans verbonden met het Nederlandse bestuur, aangezien het werd gebruikt aan het hof van de stadhouder (Frijhoff 2015: 116). Aan het eind van de zestiende en begin van de zeventiende eeuw nam het belang van het Frans als handelstaal ook steeds meer toe (Frijhoff 1989: 596; Frijhoff 2015: 116). Bovendien werd het Frans in deze tijd op steeds meer terreinen een belangrijke taal in de Nederlanden. Vanaf het einde van de zestiende eeuw vestigden zich onder andere Franssprekende vluchtelingen uit Wallonië, Frankrijk en Zwitserland in de Nederlanden, als gevolg van religieuze onderdrukking in deze landen (Frijhoff 2015: 116; Van de Haar 2019: 47). Met name de val van Antwerpen in 1585 bracht veel vluchtelingen, waaronder Franstalige Walen en tweetalige Vlaamse schoolmeesters, naar de Noordelijke Nederlanden, waardoor het taalcontact met het Frans een enorme impuls kreeg (Frijhoff 2010: 41). Deze vluchtelingen stichtten hun eigen Franstalige kerken, de Waalse kerken. Er waren er in totaal zo'n tachtig in heel Nederland. Hierdoor werd het Frans in Hollandse steden meestal de tweede taal, naast het Nederlands (Frijhoff 1989: 596; Frijhoff 2015: 116).

Sommige Franstalige vluchtelingen die zich in Amsterdam en Leiden vestigden werden uitgevers die ook Franstalige werken voor

binnen- en buitenland uitgaven, wat belangrijk was voor de verspreiding van het Frans over Europa (Burke 2004: 86–87; Frijhoff 2015: 117–118). Een voorbeeld hiervan is Jean de la Font, die naar Leiden was gevlucht en in 1680 een Franstalig politiek tijdschrift oprichtte dat bekend werd als de *Gazette de Leyde*. Andere Franstalige migranten werden leraar op zogenoemde Franse scholen, die vanaf de zestiende eeuw en in toenemende mate in de zeventiende eeuw werden opgericht in de Noordelijke Nederlanden. In 1811 waren er 147 Franse scholen verspreid over heel Nederland (Frijhoff 1996: 329). Op deze scholen, waar het Frans de instructietaal was, kwamen de leerlingen vaak uit de middenklasse; de hogere klassen gingen naar de Latijnse school (Frijhoff 2015: 119–120).

Naast een toename in het gebruik van het Frans in de media, religie en het onderwijs in de Nederlanden kwamen mannen uit de elite ook meer in aanraking met het Frans door de zogeheten *Grand Tour*, die als afsluiting van hun studie diende. Tijdens deze tour reisden ze naar verschillende landen om zich daar onder andere onder te dompelen in de taal en cultuur van het land. Tot de tweede helft van de achttiende eeuw was Frankrijk een van de landen waar naartoe werd gereisd (Frijhoff 1989: 603; Frijhoff 2015: 125–126; 136). Mede door het contact met het Frans door deze studiereis, de Franstalige migranten en de associatie van het Frans met de hofcultuur kwam er meer waardering voor het Frans en kreeg de taal een hoger prestige voor de sociale bovenlaag in de Nederlanden (Frijhoff 1989: 596; 603; Frijhoff 2015: 116; 125–126). Hierdoor werd het Frans voor een deel van de Nederlandse elite de tweede taal (Frijhoff 1990: 17).

Niet alleen in de Nederlanden, maar ook op Europees niveau kreeg het Frans steeds meer voet aan de grond en werd het vanaf de tweede helft van de zeventiende en in de achttiende eeuw de Europese lingua franca in veel domeinen (Frijhoff 1996: 329). Waar tot de zestiende eeuw het Italiaans naast het Frans als de internationale taal van handel en financiën werd gebruikt en het Latijn tot de tweede helft van de zeventiende eeuw als taal van de internationale geleterde en wetenschappelijke wereld diende, breidde het gebruik van het Frans in beide gebieden zich steeds verder uit (Burke 2004: 60; Frijhoff 2015: 118; 129; Rjéoutski & Frijhoff 2018: 7–11). Dit kwam doordat het aantal Franssprekenden vergeleken met andere Europese talen in de vroegmoderne tijd enorm was: rond 1600 waren er 14 miljoen sprekers van het Frans, tegenover 10 miljoen sprekers van het Duits, de tweede Europese taal van die tijd. Aan het eind van de achttiende eeuw waren er

zo'n 30 tot 35 miljoen Franssprekenden, tegenover 3 à 4 miljoen Nederlandssprekenden. Mede hierdoor was Frankrijk een economische en politieke grootmacht en op dat moment het machtigste land in Europa (Burke 2004: 82; Frijhoff 2010: 37; Frijhoff 2015: 128–129). Ook was het Franse hof het belangrijkste in Europa en gaf de Franse cultuur een gedragsmodel voor de elites in de rest van Europa (Frijhoff 1990: 17). Frans spreken was daarom niet alleen van praktisch nut in de internationale communicatie, maar vanwege de hoge culturele en sociale status die het Frans had, was het ook een manier om de eigen status te verhogen (Frijhoff 2010: 21–22; Frijhoff 2015: 125; 129; Rjéoutski & Frijhoff 2018: 7–11). Vanaf ongeveer 1640 is er dan ook een toenemende trend te zien waarbij de Europese elite, zoals regentenfamilies in de Nederlandse Republiek, het Frans als omgangstaal ging gebruiken (Burke 2004: 86). Geheel overstappen op het Frans hoefde echter niet als men de sociale status wilde verhogen: het gebruiken van Franse woorden of Franse frases in de eigen moedertaal werd ook gedaan in het Europa van de zeventiende en achttiende eeuw door mensen die graag wilden opklimmen op de sociale ladder (Burke 2004: 32).

In de tweede helft van de achttiende eeuw had het Frans de rol van het Latijn als *lingua franca* overgenomen en werd de taal niet meer gezien als de taal van Frankrijk, maar als internationale communicatietaal (Frijhoff 1989: 596; Frijhoff 2015: 130). Frankrijk verloor tegen het einde van de achttiende eeuw echter zijn culturele prestige langzaam maar zeker, terwijl Engeland en Duitsland een steeds belangrijker rol kregen op het gebied van onder andere literatuur, techniek en wetenschap, waardoor het Engels en Duits een grotere rol gingen spelen in de internationale communicatie (Frijhoff 2010: 38; Frijhoff 2015: 136; Argent, Rjéoutski & Offord 2014: 1). Er kan aangenomen worden dat de annexatie van de Nederlanden bij het Franse Keizerrijk aan het begin van de negentiende eeuw wel voor een nieuwe impuls heeft gezorgd voor het taalcontact met het Frans in de Nederlanden (Frijhoff 2010: 48). Rond 1850 was het Frans in Nederland in elk geval geen tweede taal meer, maar een vreemde taal, hoewel het nog wel de belangrijkste buitenlandse taal was (Frijhoff 1996: 330). Tot de twintigste eeuw bleef het Frans de belangrijkste diplomatieke taal in Europa (Burke 2004: 46).

Consequenties van Nederlands-Frans taalcontact

Het intense taalcontact met het Frans leidde tot een zekere vorm van maatschappelijke en individuele meertaligheid in het Nederlandse taalgebied (Peersman, Rutten & Vosters 2015: 1). Tweetaligheid in de strikte zin van het woord, waarbij zowel het Frans als het Nederlands goed of zelfs op moedertaalniveau werd beheerst, was een zeldzaamheid: alleen de bovenlaag van de bevolking, de hoogste ambtenaren en de succesvolste zakenmannen beheersten dit niveau in de vroegmoderne tijd (Frijhoff 2010: 53; Frijhoff 1989: 593; Van de Haar 2019: 51). Rond 1800 zou maar ongeveer vijf procent van de Nederlandse bevolking Frans hebben geleerd in een onderwijssetting (op school of van een taalleraar); na de Franse annexatie begin negentiende eeuw zal het totale aandeel van de volwassen Nederlanders dat enigszins kennis van het Frans had wellicht tien procent zijn geweest (Frijhoff 2010: 48; Frijhoff 2015: 129–130). Door het contact met het Frans zullen veel mensen iets van deze taal hebben opgepikt, maar dit zal voor de meeste vooral passief zijn geweest en op de genoemde beperkte bovenlaag na zal de kwaliteit van het Frans slecht zijn geweest (Frijhoff 2010: 46; Frijhoff 1989: 600; Van de Haar 2019: 51).

Hoe dan ook is er in de geschiedenis van het Nederlands veel contact geweest met het Frans, en dit taalcontact heeft geleid tot de overname van taalelementen uit het Frans. De precieze invloed die het Frans op de Nederlandse taal heeft gehad wordt in dit proefschrift onderzocht. Er bestaan al enkele eerdere studies naar de invloed van het Frans op de Nederlandse taal. Al begin twintigste eeuw schreef Salverda de Grave (1906) over taalelementen die het Nederlands uit het Frans zou hebben overgenomen. Ook Van der Sijs (2001; 2005) heeft overzichten gemaakt van met name morfologische en lexicale elementen die het Nederlands in de loop van de tijd uit het Frans heeft geleend. Daarnaast zijn er enkele historisch-sociolinguïstische studies naar Franse leenwoorden en leensuffixen, te weten Rutten, Vosters & Van der Wal (2015), Rutten & Vosters (2020) en Stevens (2019). Op de inhoud van deze studies wordt dieper ingegaan in hoofdstuk 5 en 7.

2.5 Onderzoeksvragen

Hoewel er door het onderzoek van Salverda de Grave (1906) en Van der Sijs (2001; 2005) bekend is wat er op het gebied van lexicon en morfologie uit het Frans is geleend, is er, afgezien van de drie genoemde

historisch-sociolinguïstische studies, weinig onderzoek gedaan naar het daadwerkelijke gebruik van deze elementen in de geschiedenis van het Nederlands. Daarnaast worden ook enkele morfosyntactische elementen in verband gebracht met invloed van het Frans, maar de daadwerkelijke relatie tussen het gebruik van deze elementen en het Frans is nooit empirisch onderzocht. Het huidige proefschrift wil hier verandering in brengen door middel van een kwantitatief corpusonderzoek naar het gebruik van Franse leensuffixen en leenwoorden in het Nederlands en het gebruik van de betreffende voornaamwoorden *dewelke* en *betwelke*, die in verband worden gebracht met invloed van het Frans, in de periode van 1500 tot 1900. Belangrijke onderzoeksvragen die hierbij worden beantwoord, zijn:

- Hoe ontwikkelt het gebruik van uit het Frans geleende suffixen en leenwoorden zich gedurende de periode 1500–1900? In welke periode tussen 1500 en 1900 werden de meeste Franse suffixen en leenwoorden gebruikt? In hoeverre komen de diachrone frequentiepatronen van de Franse suffixen en leenwoorden met elkaar overeen?
- Wat zijn de verschillen tussen sociale domeinen in het gebruik van uit het Frans geleende suffixen en leenwoorden? In welk sociaal domein komen de meeste Franse suffixen en leenwoorden voor? In hoeverre komen de frequentieverschillen tussen de domeinen voor de suffixen en de leenwoorden overeen?
- In hoeverre komt het gebruik van *dewelke* en *betwelke* overeen met de patronen die voor de leensuffixen en leenwoorden zijn gevonden? Vertoont het gebruik van deze betreffende voornaamwoorden dezelfde diachrone ontwikkeling en dezelfde frequentieverhoudingen tussen de domeinen als bij de leensuffixen en leenwoorden? Met andere woorden: in hoeverre is er bewijs te vinden dat het gebruik van *dewelke* en *betwelke* onder invloed heeft gestaan van het Frans?

Het doel van dit onderzoek is het achterhalen van de daadwerkelijke invloed die het Frans op het Nederlandse taalgebruik heeft gehad op morfologisch, lexicaal en morfosyntactisch niveau tussen 1500 en 1900. Het onderzoek wordt uitgevoerd met het *Language of Leiden*-corpus, dat speciaal voor dit project is opgebouwd. De naam van het corpus geeft al aan dat de teksten die in het corpus zijn gerepresenteerd uit Leiden afkomstig zijn. Het corpus is zodanig opgebouwd dat het representatief

is voor de historische situatie van de stad Leiden. In hoofdstuk 4 bespreek ik de corpusopbouw alsmede de redenen om het corpus op te bouwen met alleen data uit Leiden, maar eerst wordt in het volgende hoofdstuk ingegaan op de historische context van Leiden.