Towards a methodological framework for historical language choice: the case of Dutch and French in the Netherlands (1800–1899)

JILL PUTTAERT – ANDREAS KROGULL – GIJSBERT RUTTEN

ABSTRACT: This paper seeks to approach the topic of historical language choice from a quantitative perspective, arguing that solid baseline evidence drawn from a substantial dataset is a much-needed complement to the largely qualitative findings of previous research. We propose a methodological framework which enables us to examine the sociolinguistic factors that condition language choice in the private domain. Illustrating the possibilities of our methodology, we present a case study on Dutch-French language choice in the Northern Low Countries (i.e., the present-day Netherlands), focusing on nineteenth-century family correspondence. Our paper shows that a careful selection procedure is crucial in order to achieve a balanced representation of language choice in a large-scale dataset. With respect to our analyses, the role of French in private letters turns out to be relatively small against the prevalence of Dutch. However, interesting patterns become visible when looking at regional differences, gender constellations and familial relationships. These quantitative findings can therefore constitute an interpretational frame for qualitative studies on historical language choice in the Dutch-French context and beyond.

Key words: historical sociolinguistics, multilingualism, language contact, language choice, ego-documents

1. Introduction

In recent decades, the topic of language choice has attracted a fair amount of interest as part of the ever-growing body of historical sociolinguistic research on multilingualism and language contact (e.g., Rjéoutski & Frijhoff 2018). Many historical contact settings under investigation can be situated in the broader context of what is often referred to as 'European francophonie' (Rjéoutski, Argent & Offord 2014; Gretchanaia, Stroev & Viollet 2012). It describes the practice of French in areas where it was no official or vernacular language, but used as a second or foreign language. As a pan-European phenomenon, the French language spread as the major lingua franca, with its heyday traditionally located in the seventeenth and eighteenth centuries (Wright 2016: 134). It served functional needs such as diplomacy and trade, but also acquired a sociocultural dimension as a marker of distinction and a language of intimacy among the European elites.

In the Northern Low Countries, i.e., the area roughly corresponding to the present-day Netherlands, contacts with the French language go back to the early Middle Ages and continued well beyond the nineteenth century and even into the present day. This led to contact-induced changes in the Dutch lexicon and grammar, as well as to code-switching. It also led to situations of language choice, i.e., the topic of the present paper. In various domains, from trade and diplomacy, religion and education, to the private sphere, Dutch could give way to French. While the phenomenon known as *verfransing* 'Frenchification' (Frijhoff 1989), along with a strong anti-French discourse that peaked in the eighteenth century, has received a lot of attention in both historical and linguistic research

(e.g., Frijhoff 2015; Vogl 2015), empirical studies analysing the influence of French on Dutch as well as the choice between Dutch and French on a larger scale are still surprisingly scarce (Rutten, Vosters & van der Wal 2015: 146). At the same time, such studies are needed in order to test claims about the alleged Frenchification of the Northern Low Countries.

The analysis of language choice in multilingual societies is an inherently sociolinguistic issue, centred around the question famously posed by Fishman (1965: 428): "who speaks what language to whom and when". Schendl (2012: 522) points out that "[a]ll the relevant factors we need to answer this question are […] social or sociolinguistic, irrespective of whether we are investigating modern multilingual speech communities or earlier ones, though the difficulties are clearly much greater when studying earlier multilingual societies". Previous studies on historical language choice, not only in the Dutch-French context but also beyond (e.g., Offord, Rjéoutski & Argent 2018: ch. 6 on French in Russia), have typically opted for a more qualitative approach. However, we would argue that solid baseline evidence is needed in order to gain a fuller understanding of the sociolinguistic factors that determine language choice. In fact, what is still lacking in historical sociolinguistic research is a more quantitative-focused approach as "a welcome complement to qualitative analysis", which, however, can only be done with "a large enough written corpus" (Schendl 2012: 522).

As Fishman (1965: 76) already noted for 'modern' sociolinguistic studies of multilingual behaviour, "the family domain has proved to be a very crucial one". Methodologically, of course, historical situations of multilingual behaviour have to be approached differently, given the obvious focus on written rather than spoken language data. Nonetheless, the family domain plays an important role in historical sociolinguistic research, too, especially when exploring aspects of multilingualism and language contact (e.g., Head 1995; Thomas 2017; van Strien-Chardonneau 2018; van der Wal 2018). In the tradition of using handwritten archival sources from the private sphere and letter data in particular (e.g., Elspaß 2012; Rutten & van der Wal 2014; Nevalainen & Raumolin-Brunberg 2017), the present paper investigates Dutch-French language choice on the basis of a sizeable dataset of private correspondence, collected from thirty-six nineteenth-century family archives. We propose a methodological framework which enables us to examine the sociolinguistic factors that condition language choice in the private domain of family life. In fact, apart from the valuable but largely qualitative observations from previous case studies (e.g., Ruberg 2011; van Strien-Chardonneau & Kok Escalle 2017), hardly anything is known about the language choice of multilingual individuals in their everyday lives.

We first provide a sociohistorical outline of the Dutch-French contact setting and its outcomes in Section 2. Previous approaches to historical language choice are briefly discussed in Section 3. In Section 4 we introduce our newly collected data and our methodological framework. A quantitative analysis of Dutch-French language choice in nineteenth-century family letters is presented in Section 5, showcasing the possibilities of our methodology. Section 6 contains the discussion and conclusion.

2. Sociohistorical context and language contact setting

2.1. *Dutch-French contacts in the Northern Low Countries*

The Germanic-Romance language border cuts through the Low Countries, fostering language contact across the border throughout the history of the modern languages, i.e., from the early Middle Ages onwards (Peersman, Rutten & Vosters 2015). This situation of historical language contact has led to the use of both Dutch and French in a number of societal domains. In legal and administrative sources, a shift from Latin to the vernacular languages can be witnessed in the thirteenth century, when French and Dutch were increasingly used in charters (de Meyer 1974; Peersman 2012).[1] There are important regional differences, in that the use of Dutch was more widespread in the Northern Low Countries (roughly the present-day Netherlands) compared to the Southern Low Countries (presently Belgium and Luxembourg), but it should be emphasised that both vernaculars gradually replaced Latin, and subsequently developed into the two main languages of administration during the Burgundian regime in the fourteenth and fifteenth centuries (Willemyns 1994). In 1582, the States General made an explicit choice for Dutch as the political language at the highest level, though Dutch had been used for a long time already, and at the same time French remained in use, particularly in communication with non-Dutch rulers (van der Wal 1994).

French was the international language of diplomacy and trade throughout the Early and Late Modern period, though other languages were also used (Frijhoff 2015). For example, trade across the North Sea has an equally long history as trade with France, resulting in Scots and English communities in the Low Countries, and Dutch communities in England. Trade in the context of the Hanseatic League was dominated by Low German. Exploratory analyses of the sizeable collection of Early and Late Modern pamphlets kept in the Royal Library in The Hague suggests that the majority of French pamphlets published in the Netherlands were related to issues of international politics and diplomacy (Krogull, Puttaert & Rutten submitted).[2]

The Northern Low Countries, on which we focus in this paper, attracted many migrants in Early and Late Modern times, who can be roughly distinguished according to two basic types. Labor migrants were attracted by the booming economy in the Holland area in particular, and by employment opportunities with the Dutch East and West Indies Companies. A city such as Leiden in the region of Holland grew from c. 10,000 inhabitants in the late sixteenth century to 55,000 one hundred years later (Noordam 2003: 45). Many of these immigrants came from the Dutch- and French-speaking regions in the Southern Low Countries and northern France (Lucassen & de Vries 2001: 29, 40).

[1] See also Prevenier & de Hemptinne (2005) for the trilingual situation in Medieval Flanders. They provide an account of the use of Latin, French and Dutch in literary and administrative documents.

[2] The so-called Knuttel Collection, kept in the Koninklijke Bibliotheek (KB; Royal Library) in The Hague, comprises more than 30,000 pamphlets published in the Netherlands from the late fifteenth to the second half of the nineteenth century. The collection is described in Knuttel (1978), and digitally available as *Dutch Pamphlets Online* via the website of the KB.

Religious migrants included diverse groups such as Sephardim and Ashkenazim as well as Huguenot refugees from France. The latter came to the Netherlands mainly in the sixteenth and seventeenth centuries. The so-called first Refuge from the late sixteenth century onwards brought many French-speaking Protestants to the Holland and Zeeland areas (in the present-day Netherlands). The second Refuge, after the Revocation of the Edict of Nantes in 1685, brought again tens of thousands of French-speaking migrants to the Low Countries (Frijhoff 2003).

Huguenot refugees accelerated the establishment of French schools across the Netherlands. French schools were quite diverse, but gradually developed into the preferred educational track for middle and upper-middle ranked boys and, to a lesser extent, girls (Dodde 2020). The second Refuge, in particular, also brought many booksellers, printers, writers and journalists to the Netherlands, which led to a central position of the Netherlands within the international Republic of Letters (Frijhoff 2003). Cultural and intellectual life in the Low Countries had been multilingual for centuries by then. The use of Dutch as a literary language predates the thirteenth-century rise of Dutch as an administrative language (Peersman 2012). Both French and Dutch were used in medieval literature (Sleiderink 2010). There is however a clear shift towards Dutch over the centuries, particularly in the north: whereas the literary culture in the Southern Low Countries in the eighteenth century was characterised by the use of French and Dutch (Verschaffel 2017), authors in the Northern Low Countries by then almost always preferred Dutch in their literary production (Leemans & Johannes 2013).

It has traditionally been claimed that during the seventeenth and eighteenth centuries French also invaded private life, particularly among the elites (Frijhoff 1989: 596–597). Recent investigations have focused on this idea: Ruberg (2011: 69) analyses language use from five family archives from the period around 1800, and shows that 17% of their 2,302 letters were written in French. Van Strien-Chardonneau (2014) argues that French fulfilled two main functions among eighteenth-century elites, viz. as the language of international communication, and as the language of intimacy. The latter idea is taken up by van Strien-Chardonneau & Kok-Escalle (2017) who contend that French, despite being a second language to writers, was widely used in private life, for example in diaries and private letters.

2.2. *Linguistic and metalinguistic outcomes of Dutch-French contact*

Ongoing contact between Dutch and French led to several contact phenomena in both languages. Focusing on changes in Dutch, the most visible outcome of contact may be the large number of lexical borrowings. Van der Sijs (2002: 215) estimates the number of lexical items borrowed from French since the thirteenth century to be 4,605, based on an analysis of historical and etymological dictionaries. This is more than twice the number of the second largest group, viz. borrowings from Latin. Van der Wal & van Bree (2014: 173–174) argue that borrowing from French began as early as the twelfth century in trade and aristocratic contacts, and that it was quite intensive early on since the many lexical loans also led to an influx of loan suffixes in Dutch (Rutten, Vosters & van der Wal 2015;

Assendelft, Rutten & van der Wal forthcoming). Also at the level of (morpho-)syntax various changes in Dutch have been connected to influence from the contact language French, including an increase in the use of the subjunctive and participial phrases in the Early and Late Modern period (de Vooys 1970: 135; van den Toorn et al. 1997: 405; van der Horst 2008: 1150). Next to contact-induced changes, the use of French also triggered code-switching, and resulted in language choice in several contexts (see e.g., Ruberg 2011; van Strien-Chardonneau 2014; van Strien-Chardonneau & Kok-Escalle 2017).

The contact setting furthermore led to a debate on the function and use of French vis-à-vis Dutch in the Low Countries, which eventually developed into a long-term discourse of alleged Frenchification, encompassing not only language and literature but culture as a whole (Frijhoff 1989, 2015). This anti-French discourse concerned such diverse topics as the opposing political cultures in aristocratic France and the republican Netherlands (Kloek & Mijnhardt 2001: 77–78), the use of French among supposedly native Dutch speakers, and the much-needed avoidance, or even erasure of lexical borrowings from French as evidenced by the tradition of puristic dictionaries (Rutten, Vosters & van der Wal 2015).

3. Previous studies on historical language choice

Studies on historical multilingualism in the Dutch context, often in relation to French, have predominantly explored the topic of language choice from a qualitative perspective, focusing on specific cities (e.g., Kessels-van der Heijde 2002, 2015) and well-known individuals (e.g., Joby 2014) or families (e.g., van Strien-Chardonneau 2018). In the following, we first briefly discuss those studies on historical language choice in the Dutch-French context. Then, we provide a concise overview of previous studies in other contexts of historical multilingualism (i.e., Russia, Prussia and Spanish Louisiana).

Kessels-van der Heijde (2002), in her study on Dutch-French language relations in nineteenth-century Maastricht, analyses the use of Dutch, French and the Maastricht dialect in several sociocultural domains. She considers the administrative domain, the educational domain, the domain of written media (e.g., newspapers, almanacs, books), but also the familial domain, for which she draws on private correspondence and administrative documents of all kind (e.g., bills and deeds). Kessels-van der Heijde's analysis of multilingualism and language choice in the archives of nine families from Maastricht is rather descriptive, without a clear quantitative basis. Her work still provides extensive insights into the linguistic relations in nineteenth-century Maastricht, albeit from a largely qualitative perspective.

Another study on the Dutch context is conducted by Joby (2014), offering a comprehensive overview of the multilingualism of the Dutch statesman Constantijn Huygens (1596–1687), who used no less than eight languages (Dutch, French, Latin, Greek, Italian, English, Spanish, (High) German) (Joby 2014: 15). For his analysis, Joby draws on a multitude of primary sources, such as Huygens's poems and music, private correspondence to family, friends and like-minded peers, as well as letters written and received in the context of his administrative functions. Joby's work provides a profound description

of the languages used by Huygens in different contexts and domains. He first and foremost takes a qualitative approach, only occasionally providing distributional patterns, too (see e.g., Joby 2014: 98 for provisional figures on language choice in Huygens's correspondence).

While Joby focuses on one prominent historical individual, van Strien-Chardonneau (2018) analyses the practice of French in one patrician family, the van Hogendorp family, whose members held various administrative positions in the province of Holland and on a national level. Looking at three generations, she examines texts intended for publication (e.g., writings on literature, economics and politics) as well as ego-documents, or first-person accounts, such as letters, diaries, journals and autobiographical writings (van Strien-Chardonneau 2018: 67–68), discussing the different functions French served within this family (e.g., as language of distinction, and as language of self-narration and private life). Furthermore, van Strien-Chardonneau pays attention to instances of mixed-language use, such as French-Dutch code-switching, and more generally tries to determine which factors may influence the choice of language. She suggests that the gender of the writer, the political context and identity awareness play an important role (van Strien-Chardonneau 2018: 76). Although she addresses a number of interesting concepts and uses a sizeable set of primary sources, her analyses are also mainly qualitative in nature. This means that we do not gain insights into the exact distribution of French versus Dutch in the larger community.

Ruberg (2011), investigating the correspondence of several Dutch elite families between 1770 and 1850, approaches language choice from a more quantitative point of view, in addition to her qualitative analysis. She analyses the distribution of Dutch and French across a number of variables, such as the period in which the letter was written, the gender of the writer and the combination of the gender of the writer and the recipient. Although she draws on a large dataset of more than 2,300 letters, her analyses are based on a fairly small number of five families from only three different cities (Den Bosch, The Hague and Leiden). Ruberg (2011: 70) herself indicates that her corpus may not be entirely representative, since "it constitutes a mere fraction of the total actually exchanged in the eighteenth and nineteenth centuries". Moreover, the 'preservation policy' of family archives most probably conditions the availability of primary sources, too (see also Section 4.1). She therefore concludes that it is "perhaps far more revealing to approach the question of language choice from a more qualitative perspective" (see also Schendl 2012: 522).

Beyond the Dutch-French context in the Northern Low Countries, we also find hardly any quantitative studies based on larger datasets. Investigating the French language in Russia, Offord, Rjéoutski and Argent (2018: ch. 6) devote a chapter to 'writing in French'. Partly, the focus in this chapter is on the Russian literary society, but the authors also look at various ego-documents. Their analysis of language choice among the high nobility focuses on personal correspondence, diaries, travel notes, memoirs and albums produced by a select group of prominent literary figures as well as high officials working at the court. Taking a primarily qualitative approach, Offord, Rjéoutski and Argent (2018) provide a general overview of language use and choice in these documents, while also presenting instances of code-switching.

In the Prussian context, Böhm (2010) explores the linguistic acculturation of French-speaking Huguenots from the seventeenth until the nineteenth century. Comparing three colonies (Berlin, Strasburg/Uckermark, Battin), she draws on a corpus of various handwritten documents (administrative texts such as church records, as well as private and business letters) from Huguenot writers. In four separate empirical studies she discusses strategies of language choice, language shift, multilingualism and language learning in different (social) contexts and domains. Böhm's approach to language relations is also mainly qualitative.

Moving to historical studies beyond the phenomenon of 'European francophonie', Thomas (2017) studies private correspondence of the French-Spanish bilingual Bouligny family in the Late Modern period. Her corpus comprises letters written between Spanish Louisiana, France and Spain, as the family (originally from Italy and later France) emigrated to Louisiana and Spain. While she concentrates on the linguistic effects of bilingualism, using both quantitative and qualitative methods, she also analyses language choice and instances of code-switching. Predominantly from a qualitative angle, she looks at language choice within the social network of the writers under scrutiny.

This concise overview of previous research on language choice, not only in the Dutch-French context but also in other historical contexts of multilingualism, shows that sizeable empirical studies based on large datasets and quantitative analyses are still sparse, if not lacking altogether. We would argue, however, that such solid baseline evidence is essential in order to gain insights into the complex issue of language choice and the sociolinguistic factors that influence it. Therefore, we propose a methodological framework to analyse language choice and the parameters determining this choice from a more quantitative perspective.

4. Developing a methodological framework for historical language choice

In order to examine Dutch-French language choice in the Northern Low Countries, we developed a methodological framework that enables us to approach this complex phenomenon quantitatively. Complementing the (valuable) qualitative observations from previous studies (see Section 3), our central research aim is to assess "the dynamics which determine language choice in circumstances where knowledge of more than one language makes choice possible" (Offord 2020: 14), again recalling Fishman's (1965) famous question. In this paper, we investigate language choice in the social domain of private life, utilising the wealth of family correspondence kept in Dutch archives. The relatively wide availability of such letter collections particularly applies to the nineteenth century, which is the period that we focus on.

In what follows, we discuss the most important stages of our methodological approach, viz. the archival data collection (in 4.1), the compilation of databases (in 4.2), the categorisation of language choice (in 4.3), the selection of letter data (in 4.4), and the representation of language choice (in 4.5). Finally, the dataset used for this case study is outlined in 4.6.

4.1. *Archival data collection*

For the purpose of investigating Dutch-French language choice in the private domain, we collected a substantial amount of original manuscript (i.e., unedited) letter data from family archives across the Netherlands. These documents were for the most part photographed by members of our team, though occasionally provided by the archives as scans. Aiming for a wide geographical coverage of the Northern Low Countries, we selected twelve cities from ten different provinces. Listed in alphabetical order, these cities (with their respective provinces given in brackets) include: Amsterdam (North Holland), Arnhem (Gelderland), Den Bosch (North Brabant), Groningen (Groningen), Haarlem (North Holland), Leeuwarden (Friesland), Leiden (South Holland), Maastricht (Limburg), Middelburg (Zeeland), The Hague (South Holland), Utrecht (Utrecht), and Zwolle (Overijssel).[3] A map is provided in Figure 1.

Socioeconomically and demographically, the selected families belonged to the contemporary urban elite, i.e., the nobility, aristocracy, patriciate and emerging bourgeoisie living in Dutch cities. These higher and educated strata of society are traditionally associated with the use of French, also as "a sign of recognition between people belonging to the same social group" (van Strien-Chardonneau 2014: 171). With respect to the targeted coverage of urban space, we selected those families which display strong links with their respective cities, or at the very least with their respective regions or provinces. However, we have to keep in mind the geographical mobility of individual family members or even whole generations, for professional and/or personal reasons. It was not uncommon for members of this mobile urban elite to move to other parts of the country or even to places abroad.

While focusing on Dutch-French language choice, it is important to note that we did not select families based on their linguistic profiles, prioritising, for instance, Dutch family archives with a high proportion of French-language documents. Instead we selected families which, from a socioeconomic and demographic point of view, belonged to the groups most strongly associated with 'francophonie' and thus could (potentially) make a language choice for either Dutch or French in their private correspondence. More practically, our selection of families was, of course, also determined by the availability of family archives, some of which contained more comprehensive collections of private correspondence than others. Finally, within the family structure, we prioritised letter data representing the closest relations of the immediate family, viz. parents, children, siblings, and spouses. More distant relations of the extended family, such as grandparents, grandchildren, aunts and uncles, nieces and nephews, cousins, as well as in-laws, were considered when data from the immediate family proved to be limited.

[3] Data was collected from the following archives: Stadsarchief Amsterdam (Amsterdam), Gelders Archief (Arnhem), Brabants Historisch Informatie Centrum (Den Bosch), Groninger Archieven (Groningen), Noord-Hollands Archief (Haarlem), Historisch Centrum Leeuwarden (Leeuwarden), Erfgoed Leiden en Omstreken (Leiden), Regionaal Historisch Centrum Limburg (Maastricht), Zeeuws Archief (Middelburg), Haags Gemeentearchief (The Hague), Nationaal Archief (The Hague), Het Utrechts Archief (Utrecht), and Historisch Centrum Overijssel (Zwolle).

4.2. *Compilation of databases*

As described in 4.1, a substantial number of private family letters was collected for the specific purpose of assessing Dutch-French language choice in the Northern Low Countries, serving as baseline data for the study of (sociolinguistic) factors determining the choice of languages. Based on digitised, though untranscribed manuscript sources, detailed inventories or databases were compiled for all families, each comprising the same extensive set of metadata.

Each database contains archival information, such as the name of the archives from which the letters were collected, the name of the family and their associated city, the archive and inventory numbers of the letter, as well as the file names of the digitised images. Furthermore, each database includes metadata related to the letters themselves, i.e., the date and place of writing as well as the place of address (if available), which were also turned into more standardised temporal (e.g., century, or '1800–1849', '1850–1899') and spatial parameters (e.g., provinces, also including non-Dutch options like France, Belgium or 'abroad'). Crucially, metadata on the letters also covers the language choice, which we discuss in detail in 4.3 below.

Next, we retrieved as much biographical information as possible for all letter writers. These metadata include the writer's name, gender, date and place of birth, as well as date

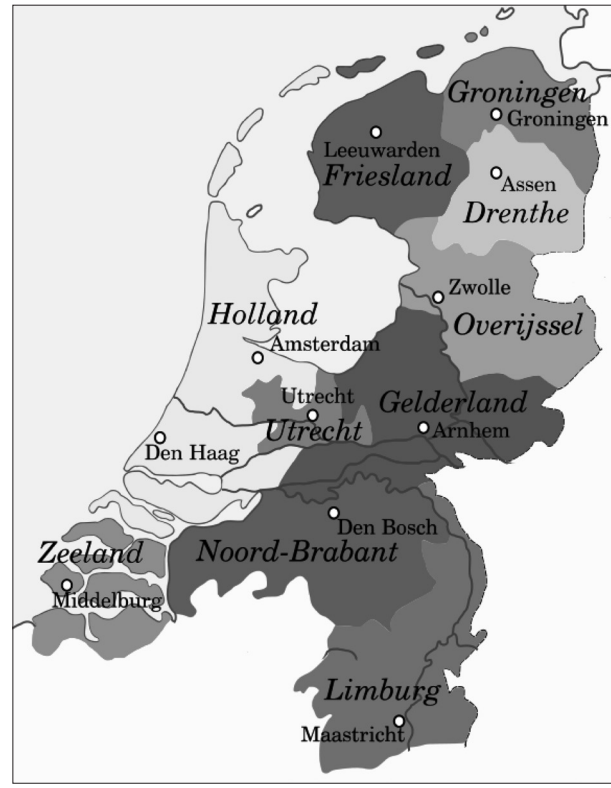

Figure 1: Map of the Northern Low Countries (as part of the United Kingdom of the Netherlands, 1815–1830)

and place of death. We also considered a range of sociolinguistic parameters determining the relationship between writers (or senders) and addressees, i.e., the communicative setting of letter writing in which a language choice is made. Here in particular, we operationalised metadata into variables that can be analysed quantitatively. These include familial relationships (i.e., child–parent, parent–child, siblings, spouses, plus more distant relationships with the extended family), gender constellations (i.e., male–male, male–female, female–male, female–female, plus constellations with 'mixed' genders), as well as the number of senders and addressees (i.e., single–single, single–multiple, multiple–single, multiple–multiple). A typical example of 'mixed' gender and 'multiple' number is a letter written by or addressed to both parents. Finally, we added the names of the addressees, as well as miscellaneous notes from code-switching examples to metalinguistic comments or the document's material condition.

4.3. *Categorisation of language choice*

Undoubtedly the most crucial parameter incorporated in our database design is 'language choice', which we had to assign to each text. The categorisation of linguistic categories was largely informed by the process of inventorying and manually reading through all letters. We opted for five linguistic categories: (i) 'Dutch', (ii) 'Dutch/French', (iii) 'French', (iv) 'French/Dutch', and (v) '50/50'. In this section, we define these five linguistic categories used to determine 'language choice' in private family correspondence. These categories are illustrated with authentic examples from our nineteenth-century dataset (see 4.6).[4]

To begin with, categories (i) 'Dutch' and (iii) 'French' refer to the two most monolingual language choices, either with Dutch or French as the main language.[5] It is important to note, though, that these two categories may not be entirely monolingual in the strictest sense. In fact, to a certain extent we allowed the presence of other-language elements such as borrowings,[6] single-word switches (see Fr. *franchement* 'frankly' (1); Du. *Wiskunde* 'mathematics' (2)) or short multi-word switches functioning as lexical units (see Fr. *de bon cœur* 'willingly' (1); Du. *huwelijkse voorwaarde* 'prenuptial agreement' in (3)). Note that other-language elements are sometimes marked in the original text, for instance by underlining as in (2) and (3).

[4] All examples were transcribed diplomatically from the original manuscripts, including the use of accent marks. We highlighted the other-language elements in bold.

[5] We prefer to avoid the controversial terms *matrix* and *embedded language* introduced by Myers-Scotton (1993). Her Matrix Language Frame model tends to be too rigid when studying written practices in historical settings of multilingualism (see also Pahta, Skaffari & Wright 2018: 8). As Gardner-Chloros (2018: 24–25) critically puts it, 'it is unclear what is to be gained, apart from theoretical tidiness, by describing multilingual texts as being made up of a "matrix" language (actually a matrix grammar) and its complement, an "embedded" language'.

[6] We are aware of the debates regarding conceptual differences between borrowings and (single-word) code-switching (see e.g., Stell 2019 for a concise overview). However, since this (fuzzy) distinction is not relevant for our linguistic categorisation, we will not delve into this discussion at this point.

(1) *Wat uw raad betreft dat Sophie, als gij het ouderlijk dak gaat verlaten, eene jufvrouw zou nemen, moet ik **franchement** afkeuren, tenzij pa er zich geheel **de bon coeur** bij neerlegt: doch daarover mondeling nader.*

'As regards your advice that Sophie, when you leave the parental roof, should take a maid, I **frankly** have to reject it, unless Dad **willingly** accepts it: but more about that later verbally.'

(2) *Ce matin nous avons eu **Wiskunde**. Cela n'a pas été très bien car déjà au commencement il y avait une somme que je ne savais pas très bien alors je me suis tellement agité que je pleurais d'agitation.*

'This morning we had **mathematics**. It didn't go very well because already at the beginning there was a sum that I didn't know very well, so I got so agitated that I cried with agitation.'

(3) *vous sentez, que cette ignorance même, s'avoir de qu'elle maniere vous avez fait avec votre Epouse le **huwelykssche voorwaarde** me met dans l'impossibilité de vous donner des conseils soit a l'egard de vos finances, soit a l'egard de votre plan de separation*

'you feel that this very ignorance, knowing in what way you have done the **prenuptial agreement** with your wife, makes it impossible for me to give you advice, either with regard to your finances, or with regard to your separation plans

Further types of other-language elements included in categories (i) and (iii) comprise titles of, for instance, books and theatre plays, as shown in examples (4) and (5), respectively:

(4) *Ik ben aan een lecture over het van Kant maken bezig. maar die manier is veel zachter en doet geen pyn. t' is in t' fransch geschreven en heet **"Histoire de la Dentelle"***

'I am currently reading about lace-making. But this manner is much softer and does not hurt. It is written in France and is called **"Histoire de la Dentelle"** [= History of Lace]'

(5) *Er zal komedie gespeeld worden **le Medecin malgré lui** van Molière, en natuurlijk gedanst worden.*

'There will be comedy **Le Médecin malgré lui** [= The Doctor in Spite of Himself] by Molière, and of course dancing.'

We also allowed for other-language proper nouns of persons and places, as well as other-language dates in our two 'monolingual' categories. In (6), the place and date are written in French (with the exception of Du. *uur* 'o'clock'), whereas the main body from the salutation onwards is written in Dutch.[7]

(6) ***Bruxelles, Hôtel de l'Univers***
8 Mars 1856
***Samedi soir**, 11 ½ uur*
Beste Ouders,
Ik profiteer van een rustig half uurtje, dat wij na onzen komediegang van heden avond hebben, om U te vertellen hoe goed & hoe aangenaam wij het tot nog toe gehad hebben.
'**Brussels, Hôtel de l'Univers**
8 March 1856
Saturday evening, 11 ½ o'clock

[7] Like Thomas (2017: 87), we do not consider the language choice of the address to be "part of the same communicative event because the intended audience is not the addressee, but rather the postal service or letter carrier". What is more, the address (either on the outside of a letter or on the envelope) is not necessarily available for all texts, which makes it difficult to consistently apply this additional level to a larger dataset.

Dear parents,
I am taking advantage of the quiet half hour that we have after our visit to the comedy tonight, to tell you how good and how pleasant we have had it so far.'

Similar to categories (i) and (iii), there is one dominant language in categories (ii) 'Dutch/French' and (iv) 'French/Dutch'. In category (ii) the main language is Dutch, whereas it is French in category (iv). However, the presence of other-language elements is more substantial than in (i) and (iii), to the extent that code-switching becomes characteristic of the language choice in these two categories. For instance, code-switching may occur as strings of (inserted) multi-word switches or as (alternating) clauses, sentences or entire passages in the other language, both within the 'main body' or as a postscript at the end of a letter. As a rule of thumb, a minimum of three lines in a text containing multi-word switches into the other language was required in order to be assigned to categories (ii) or (iv). In examples (7)–(14), we illustrate these different types of code-switching from Dutch to French and from French to Dutch, which all occur in letters falling into categories (ii) and (iv), respectively. These comprise inter-sentential switches as in (7) and (9), as well as intra-sentential switches as in (8) and (10). Note that in handwritten archival sources in particular, the boundaries between intra- and inter-sentential code-switching can be blurry at times (see e.g., (11)).

(7) *je veux cependant saisir cette occasion pour vous écrire qques lignes, & vous assurer que nous sommes bien portants, & tres disposes de nous bien amuser à Paris; apres demain nous nous mettons en route. J'espere d'y arriver sain & sauf,* ***het is dog een onderneming om er dit saisoen nagt & dag in de dilligence te zitten***

'I would like to take this opportunity to write you a few lines and assure you that we are in good health, and very prepared to enjoy ourselves in Paris. The day after tomorrow we are setting off. I hope to arrive there safe and sound, **it is quite an enterprise to sit in the coach night and day this season'**

(8) *maar hebt gij de zwarte oude Pantalon van Charles, ik bedoel die door Telman gemaakte, bij abuis ook mede genomen? Ze is hier in huis nergens meer te vinden. Alom* ***me je méfie ma servante, Je vous prie de me repondre à ce sujet par le retour du courier***

'But did you also take Charles's old black trousers, I mean the ones made by Telman, with you by mistake? They are nowhere to be found in the house. **I distrust my maid** all over, **I beg you to answer me on this subject by return of the mail**'

(9) *Je viens te donner qq mots de nos nouvelles mais ce ne pourra être une longue lettre car je souffre d'un érésilpèle sur les yeux qui m'incommode et m'empêche de bien voir* ***ik heb een lap op het oog hangen als een Koe in de Weide****.*

'I come to send you a few words on our news, but it cannot be a long letter because I am suffering from erysipelas on my eyes, which bothers me and prevents me from seeing well. **I have a patch on my eye like a cow in the meadow**.'

(10) *vous voila bien contente j'en suis sure c'est un grand amusement pour vous car je crois que vous aurez souvent la petitte sur le bras et je suis sur que Henriette* ***er zig niet scheef aan draagen zal*** *car elle me ressemble elle n'aime pas* ***van die kleine gevilde konyntjes****.*

'you will be very happy, I am sure thereof, it is a lot of fun for you because I believe you will often have the little one on your arm, and I am sure Henriette **will not carry it that much** because it seems to me that she does not like **those little skinned rabbits**.'

(11) ***Daar zull[en] wij naar kerk gaan, en de*** *la servante dit à nous que c'est encore 8 heure (à 9 heure elle commence). Nous avons resté si longtemps que la même servante nous dit que c'était 9 heure, nous allons, mais c'était deja 9 et demi*

'**There we will go to church, and the** the maid said to us that it was still 8 o'clock (at 9 o'clock it starts). We stayed so long that the same maid told us that it was 9 o'clock, we went, but it was already 9.30'

Multi-word switches in categories (ii) and (iv) do not necessarily have to be 'creative' text as in (7)–(11), but they may also contain proverbs, quotations or copied passages from other texts. For instance, (12) illustrates a switch from Dutch to French, in which the writer cites a character from Molière's comedy *Les Femmes Savantes*. In (13), we find a French proverb, loosely referencing philosopher Voltaire. In (14), the writer copies a passage in French from a local newspaper (*Maastrichter Courant*) into his otherwise Dutch letter.

(12) *Ik zal maar zeggen zoo als Martine de meid in les femmes savantes* ***"quand on se fait entendre, on parle toujours bien, Et tous les biaux dictons ne servent pas de rien."***

'I'll just say like Martine the maid in Les Femmes Savantes [= The Learned Ladies] **"when we are understood, we always speak well, and then all your fine diction serves no purpose."**'

(13) *Ik besprak met hem verder de huishoudelyke regeling: aanschaffing van collegestoelen en tafels en van een zwart bord via de Boer.* ***Enfin tout pour le mieux dans le meilleur des mondes****.*

'I further discussed with him the household arrangement: purchase of lecture chairs and tables and a black board through De Boer. **In the end, all for the best in the best of worlds**.'

(14) *zie hier alles wat ik weet over de doorbraak van het Cannaal te Neeroeteren gecopieerd uit de Maastrichter Courant –* ***Nous apprenons qu'avant hier 16 la digue droite du Zuid Willems Vaart à été Rompue a Neeroeteren Canton de Maeseyck à environ 50 pas au nord du pont*** **[…]**

'See here all that I know about the breakthrough of the Canal at Neeroeteren copied from the Maastricht Courant – **We learn that the day before yesterday the 16th the right dyke of Zuid-Willemsvaart was breached in Neeroeteren, canton of Maaseik, at about 50 paces north of the bridge […]**'

Finally, we added category (v) '50/50' to account for those texts in which Dutch and French are used to more or less the same extent. These instances can range from texts with neatly separated parts in Dutch and French, with the first half being written in one language, and the second half being written in the other language, to cases of constant code-switching between Dutch and French, both inter- and intra-sentential. The latter is illustrated in (15), where we highlighted the Dutch-language parts in bold.

(15) *Je suis donc ici depuis mercredi passé le 17 du courant quand a eu loué mon installation* ***zooals gij begrypt met zeer veel deftigheid en staatsie.*** *Pour le moment je suis encore logé à Bingerden mais la semaine prochaine je vais prendre mon* ***intrek*** *à l'hôtel* ***"Het wapen van Bingerden"*** *où j'ai loué une chambre qui fonctionnera comme salon, salle à diner, chambre à coucher, etc etc.* ***Ik heb my daar besteed voor 25 stuivers per dag behalve vuur en licht. Gij begrypt dus dat daarbij niet veel vet in de pot kan komen en ik wel mager zal worden. In het voorjaar denk ik een kasteel te bouwen echter niet om er mijn intrek te nemen met vrouw en kinders.*** *Mon depart du Zijp m'a coûté bien de la peine car vous me connaissez le grand faible que j'avais pour cette cherie terre natale mais je comprenais qu'on m'en chasserait pourtant tôt ou tard*

et mieux vaut partir libre que d'être chassé. ***Ik zal dus nu maar moeten zien*** *tant bien que mal* ***myn eigen huishouden op te rigten****. Au Zyp pour le reste* ***blijft alles bij het oude***

'I have been here since last Wednesday the 17th, when I rented my residence, **as you understand with very much pomp and splendour**. For the moment I am still staying in Bingerden but next week I will **move** into hotel **"Het wapen van Bingerden"** (= The arms of Bingerden) where I have rented a room that will function as living room, dining room, bedroom, etc. etc. **I have taken up residence there for 25 stuivers (= five-cent coins) there, excluding fire and light. You understand that I will not save a lot of money and I will become skinny. In spring I think I will build a castle, but not to live there with my wife and children.** My departure from Zijp cost me a lot of pain because you know the great weakness I had for this dear home land, but I understand that I would be chased away from it sooner or later and it is better to leave freely than to be chased away. **I will thus have to see now that I set up my own household** as best I can. At Zijp everything **stays the same** for the rest'

Whatever form these fifty-fifty cases take, they render the identification of one single dominant language impossible (let alone one grammatical frame in the sense of a matrix language) and therefore justify a linguistic category in its own right.

One could argue that even a categorisation of five 'language choices' in the Dutch-French context is a simplified representation of a much more diverse linguistic reality. One may think of further distinctions between types of code-switching or the presence of code-switching to languages other than Dutch or French (which we deliberately excluded here). Nonetheless, given our quantitative-oriented research aims, these five categories allow for a fine-grained account of language choice that moves beyond the traditional categorisation of language X versus language Y and perhaps one mixed-language category (comprising languages X and Y) (e.g., Ruberg 2011: 259–261).

4.4. *Selection of letter data*

With respect to the representative selection of private family correspondence, a well-defined set of selection criteria seemed crucial at three levels: families, writers, and sender–addressee relationships. First of all, we have to consider major differences regarding the scopes of family archives, which can range from copious to more fragmentary collections of letters. Secondly, at the level of individual writers, we usually find a few (often male) family members being overrepresented in the preserved correspondence (cf. Ruberg 2011: 70 on 'preservation policy'). Thirdly, at the level of sender–addressee relationships, archives may have preserved dozens or even hundreds of letters by a certain writer addressed to, say, their father, but only a handful of the same writer to their sister. In this example, the writer's language choices made in letters to their father are overrepresented in absolute terms (compared to those made to their sister). We therefore advocate a careful selection procedure in order to establish a balance in such an unevenly distributed mass of archival data, both within and across family archives. The selection criteria for all three levels are outlined below.

Taking into consideration the divergent scopes of family archives, we defined a maximum of forty letters per family, levelling the differences between larger and smaller collections of private correspondence. For this selection of (up to) forty letters per family,

we draw on the entirety of inventoried letters in the databases compiled for all family archives (see 4.2 above). Selection at the levels of the individual writer and the sender–addressee relationship is interconnected. When possible, we selected texts by at least five different letter writers of each family. Furthermore, at the level of sender–addressee relationships, we defined a maximum number of four different addressees per writer. We also defined a limit of three letters per sender to the same addressee. This comes down to a maximum number of twelve texts of the same letter writer in the final data selection (distributed over up to four sender–addressee relationships), preventing an overrepresentation of language choices made by particularly prolific writers.

4.5. *Representation of language choice*

It is this very level of the sender–addressee relationship at which we represented the language choices of the (letter) writer (see 4.3 for our five linguistic categories). In other words, we aimed to represent language choice in the unique relationship between one letter writer (or sender) and one specific family member (or addressee),[8] rather than the individual's outgoing family correspondence in its entirety. This essential difference in approaching language choice is also emphasised by Fishman (1965: 76), who argues that an approach considering the *dyadic* relationships within the family not only recognises that "interacting members of a family […] are *hearers* as well as *speakers* (i.e., that there may be a distinction between multilingual *comprehension* and multilingual *production*)", but also that "their language behaviour may be more than merely a matter of individual preference or facility but also a matter of *role-relations*" (see also Head 1995: 592).

We give three simplified examples in order to illustrate the important methodological stage of representing Dutch-French language choice in private family correspondence. First, if all inventoried letters within a sender–addressee unit are written in the same linguistic category, for instance 'Dutch', then three letters are selected to represent the language choice within this unit.[9] Secondly, if there are two linguistic categories in the inventoried letters of a sender–addressee unit, of which a majority is written, say, in 'French', with some additional letters written in 'Dutch', then the selection of three letters comprises two texts representing the dominant language choice ('French'), and one text representing the other language choice ('Dutch'). Thirdly, if we encounter three different linguistic categories in the inventoried letters of a sender–addressee unit, each of these language choices is represented by one text for each category, irrespective of their exact proportion. Note that we allowed for two exceptions to these selection criteria. As a first exception, we may encounter an equal number of texts written in, for instance, 'Dutch' and 'French', within the same sender–addressee unit. In this case, only two (rather than three) texts are

[8] In a similar vein, Thomas (2017: 87) takes "the bounds of a letter, defined as the same writer(s) to the same addressee(s), as the communicative event".

[9] Our selection at this stage was relatively random. Priority was given to letters with the most metadata, though, especially the date and place of writing. Readability of the handwriting was another, more secondary factor.

selected, each of which represent a linguistic category. As a second exception, the maximum number of three letters within the same sender–addressee relationship can be exceeded if more than three different linguistic categories (i.e., four or five) occur. Importantly, the maximum number of letters per writer remains twelve nonetheless.

4.6. *Dataset*

For the quantitative analysis in Section 5, which showcases the opportunities of our methodological framework introduced above, we focus on private family correspondence from the nineteenth century, i.e., letters written in the period between 1800 and 1899. This dataset consists of letters collected from thirty-six Dutch family archives from across the language area. As shown in Table 1, the dataset covers twelve cities from ten provinces in the Northern Low Countries. Each city is represented by three families. Our nineteenth-century dataset contains 1,329 private family letters (which is a representative selection from more than 7,000 inventoried texts in total). The selected texts were produced by 371 individual letter writers, equally distributed across men and women (both 50%). With respect to the level of representing language choice, i.e., our sender–addressee unit, the dataset comprises 563 unique sender-addressee relationships. Whenever possible, we sought to include genders (for both senders and writers), familial relationships and generations in a balanced manner.

Table 1: Dataset of nineteenth-century private family correspondence

City (Province)	N families	N letters	N writers	N sender–addressee units
Amsterdam (North Holland)	3	119	33	54
Arnhem (Gelderland)	3	105	33	45
Den Bosch (North Brabant)	3	120	27	50
Groningen (Groningen)	3	120	45	59
Haarlem (North Holland)	3	101	26	41
Leeuwarden (Friesland)	3	103	24	41
Leiden (South Holland)	3	120	34	50
Maastricht (Limburg)	3	96	29	47
Middelburg (Zeeland)	3	92	30	37
Utrecht (Utrecht)	3	120	33	45
The Hague (South Holland)	3	113	25	42
Zwolle (Overijssel)	3	120	32	52
Total	**36**	**1,329**	**371**	**563**

5. Quantitative analysis

In this section we apply the quantitative method presented in Section 4, illustrating its possibilities by focusing on a range of variables that may condition language choice, viz. cities and regions (in 5.1), gender and gender constellations (in 5.2), and familial relationships (in 5.3).

Looking at the overall distribution of language choice in the nineteenth-century dataset, we see a strong prevalence of Dutch (Table 2). The share of the 'Dutch' category accounts for 72%, while the 'French' category constitutes only 19.6%. The two mixed-language categories are considerably smaller, but equally distributed, with 'Dutch/French' and 'French/Dutch' accounting for 3.4% and 3.3%, respectively. Finally, the '50/50' category constitutes 1.7%. It is clear that the position of French in this dataset should not be overestimated, although it is not a marginal phenomenon either, since letter writers sometimes opted for French (i.e., in 327 out of 1,329 letters in total, if we add up the three categories with a considerable share of French, viz. 'French', 'French/Dutch' and '50/50'). Moreover, the fact that Dutch is the preferred language in our nineteenth-century data should come as no surprise, since the heyday of French influence in the Northern Low Countries is traditionally situated in the seventeenth and eighteenth centuries (Argent, Rjéoutski & Offord 2014: 1; van der Wal & van Bree 2014: 254; cf. also Wright 2016: 134).

Table 2: Relative distribution of language choice in the nineteenth-century dataset (N = 1,329)

Dutch		**Dutch/French**		**French**		**French/Dutch**		**50/50**	
N	**%**	**N**	**%**	**N**	**%**	**N**	**%**	**N**	**%**
957	72.0	45	3.4	261	19.6	44	3.3	22	1.7

5.1. *Cities and regions*

As mentioned in 4.1, our dataset aims for a wide geographical coverage and thus includes families from twelve cities from different provinces of the Northern Low Countries (see Figure 1 for a map). In this section, we focus on regional differences and discuss the possible influence of the variables of 'city' and 'region' on the choice of language.

If we look at the distribution of language choice across the twelve cities under scrutiny (Table 3), it becomes apparent that French is used in every city. The choice of French is thus not confined to those cities where one might expect it, such as the capital Amsterdam, the court capital The Hague, or cities closer to the Germanic-Romance language border (such as Maastricht) and, more generally, cities near the Southern Low Countries, which are therefore relatively close to the language border (Den Bosch, Middelburg). Even in the far north (Leeuwarden, Groningen), French can be attested, too. However, it should be stressed that the percentages across cities vary greatly, with outliers such as Maastricht, with a percentage of no less than 44% for French on the one hand, and Zwolle, where French occurs in only 6.7% of all cases, on the other. These (sometimes) large discrepancies between cities can probably be explained by the influence of individual families, as striking differences across families can be observed. In Maastricht (Limburg), for instance, the Behr family can be considered a very 'French' family, using French in more than 90% of all cases, while the correspondence of the two other families from Maastricht displays a clear preference for Dutch. The same applies to Haarlem (North Holland), where the van Styrum family stands out as another predominantly 'French' family, in contrast to the two other families who indeed use French, too, but to a far lesser extent. Thus, even with a well-considered selection of three families per city, the influence of

a more atypical distribution of language choice in one family is still visible in the results and cannot be ruled out completely. This strongly suggests that analyses of language choice based on only one family per city (cf. Ruberg 2011) can hardly be representative.

Table 3: Language choice across cities

	Dutch	Dutch/French	French	French/Dutch	50/50	N
Amsterdam	74.8%	4.2%	14.3%	5.0%	1.7%	119
Arnhem	61.9%	3.8%	24.8%	6.7%	2.9%	105
Den Bosch	65.0%	3.3%	26.7%	3.3%	1.7%	120
Groningen	89.2%	0.8%	10.0%	0.0%	0.0%	120
Haarlem	39.6%	4.0%	51.5%	5.0%	0.0%	101
Leeuwarden	76.7%	3.9%	12.6%	1.9%	4.9%	103
Leiden	80.8%	5.0%	8.3%	5.0%	0.8%	120
Maastricht	52.1%	2.1%	44.8%	1.0%	0.0%	96
Middelburg	82.6%	3.3%	8.7%	3.3%	2.2%	92
The Hague	61.9%	8.0%	21.2%	4.4%	4.4%	113
Utrecht	81.7%	2.5%	13.3%	1.7%	0.8%	120
Zwolle	90.0%	0.0%	6.7%	2.5%	0.8%	120

In what follows, we look at regional differences by grouping the twelve cities into four regions: 'North' (Groningen, Leeuwarden), 'Central/East' (Arnhem, Utrecht, Zwolle), 'South' (Den Bosch, Maastricht, Middelburg), and 'West' (i.e., Holland: Amsterdam, Leiden, Haarlem, The Hague). In order to group these cities, we divided the Northern Low Countries geographically into three regions (north – central – south), based on their distance or proximity to the French language area. In the 'central layer' of the country, however, the Holland area holds a special position as the socioeconomic and demographic centre of the country, with Amsterdam being the capital city and The Hague being the city of the court (Kloek & Mijnhardt 2001), which is why Holland is best treated separately.

Turning to the distribution of language choice across region (Figure 2), interesting patterns emerge. Dutch has a particularly strong position in the north and the central/east of the country, with a share of 83.4% and 78.6%, respectively. The position of French is relatively marginal, also compared to the overall share in the country (Table 2), with percentages of 11.2% in the northern and 14.5% in the central/eastern cities. Given the distance to the Germanic-Romance language border, this is not entirely unexpected. For the northern cities, the distance to the socioeconomic and demographic centre of the Northern Low Countries (i.e., Holland) might explain the strong position of Dutch, too. French is mainly used in the south (26.9%) and the west (22.7%) of the country, although it is important to emphasise that Dutch is the prevalent language choice there as well. The relatively high preference for French in the west (i.e., Holland), in comparison with the northern and central/east parts of the Northern Low Countries, seems to suggest that the use of French can be linked to the urban elite in this (highly urbanised) Holland region, many of which held administrative positions on a supraregional and national level, where

French played a major role (see also Ruberg 2011: 69). The relatively high percentage of French in the south can possibly be explained by the region's proximity to the border with the Southern Low Countries and the French language area. Our findings echo similar explanations suggested by Ruberg (2011: 69) for Den Bosch and Kessels-van der Heijde (2002: 247) for Maastricht.

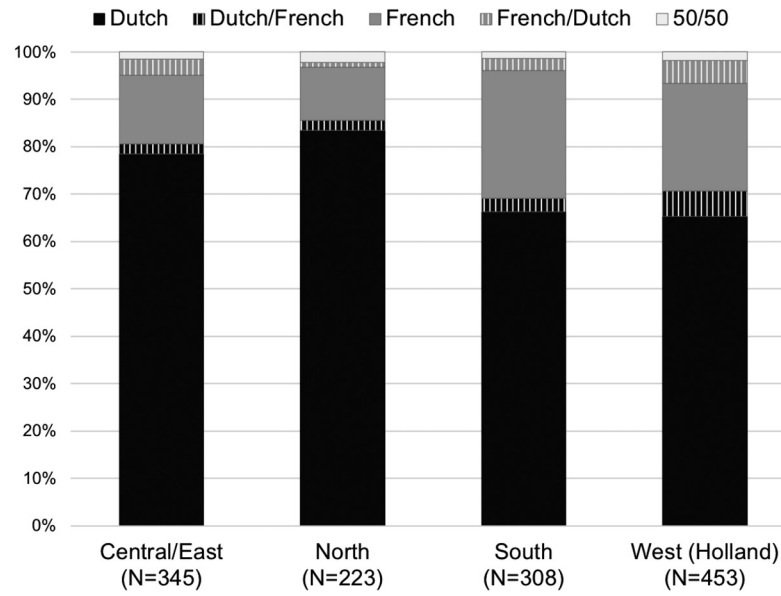

Figure 2: Language choice across region

5.2. *Gender and gender constellations*

Another sociolinguistic variable incorporated into our methodological framework is gender, by which we mean the gender of the writer. Looking at the distribution of language choice across male and female writers, no major differences can be attested (Table 4). While men opt for French in 17.9% of all cases (as opposed to 73.2% for Dutch), women use it slightly more often (22.1%, as opposed to 70.3% for Dutch), but, as mentioned, these differences are small.

Table 4: Language choice across gender

	Dutch	**Dutch/French**	**French**	**French/Dutch**	**50/50**	**N**
Men	73.2%	4.3%	17.9%	2.7%	1.9%	783
Women	70.3%	2.0%	22.1%	4.3%	1.3%	539

These results may seem unexpected with respect to previous assumptions about French being a 'women's language' (cf. van Strien-Chardonneau 2018: 76; Ruberg 2011: 70). However, if we explore gender further and consider gender constellations, taking into

account the role relations (or sender-addressee relationships) within the letters, which is essential to our methodology, a gender effect can indeed be attested (Figure 3).

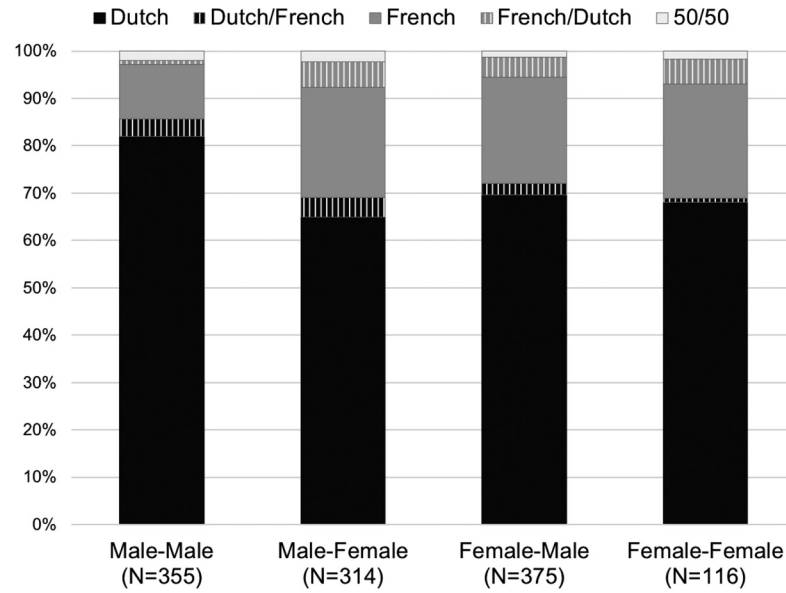

Figure 3: Language choice across gender constellations

Dutch occurs most frequently in letters written by and addressed to men, where the share of Dutch accounts for no less than 82%, as opposed to 11.5% for French. The difference with the other gender constellations (i.e., 'male–female', 'female–male', 'female–female') is striking. In letters written to and by women, the share of French is remarkably higher: men writing to women choose French in 23.2% of all cases, women writing to men in 22.4% and women writing to women in 24.1%. In particular, the differences between men writing to each other and women writing to each other are very revealing: while French occurs the least in correspondence between men, French is used most often in letters between women. These results suggest that French is more frequently used when a woman is part of the communicative setting. In other words, if a woman is involved (either as sender or addressee), we find a higher proportion of French. Ruberg (2011: 70), in her study on Dutch elite correspondence from the eighteenth and nineteenth centuries, comes to a similar conclusion, although her findings are somewhat different from ours, especially when it comes to women writing to men. In the correspondence analysed by Ruberg, women writing to men use relatively little French in comparison to the women in our dataset. Nevertheless, it is clear that the role of women within the sender–addressee relation is important with respect to the choice in favour of French. The operationalisation of gender and gender constellations into variables in our dataset thus allows for a more nuanced analysis of these variables, which proves to be a fruitful method to gain more insights into the gender dimension of language choice.

5.3. *Familial relationships*

The metadata about the relationship between senders and addressees also allow us to investigate language choice at the level of close family members (Figure 4). The most remarkable result is the difference between children writing to their parents on the one hand, and parents writing to their children on the other. We must stress, however, that we intend to analyse language choice across familial relationships irrespective of the (additional) sociolinguistic factor of age. This means that 'children' in this dataset may be ten-year-olds writing to their parents, but also, for instance, thirty-year-old adults writing to their parents. The share of French in the correspondence from children to their parents accounts for 25%, whereas parents writing to their children use French only in 10.9% of all cases. It is possible that the differences between child-to-parent and parent-to-child constellations can be explained by the hierarchical relationship among the two: children writing to their parents write 'upwards' and this might trigger the choice of French (cf. also Head 1995: 582). The results for siblings and spouses are somewhat less pronounced. French is used in 20.6% of the letters between siblings and in 22.3% of the letters between spouses.

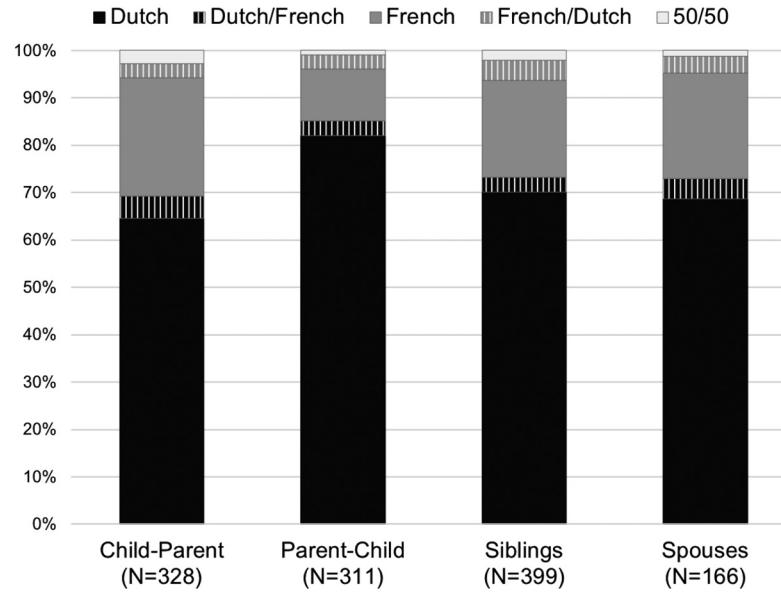

Figure 4: Language choice across familial relationships

The operationalisation of metadata about senders and addressees into variables also makes it possible to focus on the language choice across family relationships combined with gender of both senders and addressees (Table 5). Given the remarkable differences in the correspondence between children and their parents, we concentrate on these role relations, or, in fact, sender-addressee relationships. Looking at the letters written by sons

to their fathers on the one hand, and to their mothers at the other, it is remarkable that completely different language choices are made. While the two French categories combined (i.e., 'French' and 'French/Dutch') account for 19.6% in the correspondence from sons to their fathers, they constitute no less than 32.9% in the letters from sons to their mothers. Overall, the share of French in letters from daughters to their parents is bigger. If we add up the two categories that involve French as the main language, we see that daughters writing to their fathers use French in 29.3% of their letters, while they use French slightly more often in letters to their mothers (32%). French is thus more common in the correspondence with mothers than with fathers and this applies to both constellations where sons and daughters are the senders.[10]

Looking at the language choice in the letters written by parents to their children, gender seems to be a decisive factor, too. French is rarely used in the correspondence from fathers to their children: in letters written by fathers to their sons, the two categories with French only account for 8%, and for 8.6% in letters to their daughters. The contrast with language choice in the mothers' correspondence is striking: the share of the two French categories combined in letters from mothers to their sons is 21.3%, and 17.2% in letters to their daughters.

Table 5: Language choice within child-parent and parent-child correspondence

	Dutch	**Dutch/French**	**French**	**French/Dutch**	**50/50**
Child–Parent					
Son–Father (N=97)	73.2%	4.1%	19.6%	0.0%	3.1%
Son–Mother (N=79)	57.0%	5.1%	22.8%	10.1%	5.1%
Daughter–Father (N=41)	68.3%	2.4%	29.3%	0.0%	0.0%
Daughter–Mother (N=25)	64.0%	0.0%	28.0%	4.0%	4.0%
Parent–Child					
Father–Son (N=113)	87.6%	2.7%	6.2%	1.8%	1.8%
Father–Daughter (N=35)	88.6%	2.9%	8.6%	0.0%	0.0%
Mother–Son (N=108)	75.0%	2.8%	16.7%	4.6%	0.9%
Mother–Daughter (N=29)	79.3%	3.4%	10.3%	6.9%	0.0%

The quantitative analysis of a number of (sociolinguistic) variables presented in this section has shown the potential of our methodological framework, which also allows for a fine-grained analysis due to the dataset specifically compiled for this study. Some results can be discussed in even more detail (e.g., the language choice between siblings and spouses, based on the gender of the senders and addressees), which we address in future publications.

[10] The nineteenth-century dataset also includes letters from children to both their parents, but in this paper we only focus on letters written to fathers and mothers separately.

6. Discussion and conclusion

Language contact with French has been part of the history of Dutch in the Low Countries since the early Middle Ages. French was used in various domains and served different functions, from the lingua franca of international trade and diplomacy, to the language of distinction and intimacy in the private sphere, particularly among the elite. Over the centuries, encounters with the French language have led to contact-induced changes in the Dutch lexicon and grammar, to code-switching between Dutch and French, as well as to situations and issues of language choice. While the supposed 'Frenchification' of the Low Countries has received a fair amount of scholarly interest, the influence of Dutch-French contacts on language use and language choice has hardly been examined empirically.

In this paper, we investigated historical language choice in the private domain on the basis of a substantial dataset of family correspondence, collected from numerous Dutch archives. More specifically, we looked at a range of sociolinguistic factors that could determine the choice for Dutch and/or French. Historical and sociolinguistic studies on the topic of language choice in the Low Countries (and beyond) have typically taken a largely qualitative approach. In order to complement insights from previous research, we sought to establish baseline evidence on Dutch-French language choice as the interpretational frame for qualitative studies as well as for claims about Frenchification.

We developed a methodological framework that allows us to study historical language choice quantitatively. While aiming for a large enough dataset that represents the nineteenth-century urban elite, a careful selection procedure appeared to be of particular importance in order to establish a balance across and within letter collections. Family archives tend to be highly divergent in terms of the number of letters preserved, and in terms of the number and distribution of senders and addressees involved. In our family databases, we collected an extensive set of metadata, some of which were then operationalised into sociolinguistic variables, including cities and regions, gender and gender constellations, familial relationships. Methodologically, we considered the (dyadic) sender-addressee relationships between letter writers and their addressed family members to be crucial for a balanced representation of language choice. Here, we opted for a relatively fine-grained linguistic categorisation into two 'monolingual' and three mixed-language categories.

Our quantitative analyses based on a dataset of more than 1,300 selected letters have shown that Dutch was the preferred language in nineteenth-century family correspondence (72%). While the use of French was by no means marginal with a share of almost 20%, we certainly cannot conclude from these findings that this specific domain was overly Frenchified. Looking at a number of possible variables influencing language choice, the factor of region appeared to be decisive. Although French was used across the country, French was clearly more present in cities associated with the socioeconomic and demographic centre (Amsterdam, Haarlem, Leiden, The Hague) and in cities closer to the border with the Southern Low Countries and the French language area (Den Bosch, Maastricht, Middelburg).

Turning to the factor of gender (i.e., of the writer), no major differences were found between men and women. A gender effect could be attested, though, when we investigated gender constellations: French was more frequently used when a woman was part of the communicative setting (either as sender or addressee). Finally, the metadata about sender-addressee relationships enabled us to analyse language choice at the level of familial relationships. The differences between letters from children to their parents and parents to their children were particularly revealing, the former showing a considerably higher share of French than the latter. These results were even more remarkable when we combined these familial relationships with gender. Sons writing to their mothers used French more often than when writing to their fathers. Within parent-child constellations, French is barely used in letters from fathers to their children, whereas its share is considerably higher in letters from mothers to their children. The analyses of sociolinguistic parameters incorporated into our methodology have thus shown that a quantitative approach based on a sizeable dataset can be very fruitful to gain insights into the sociolinguistic dynamics of language choice.

Such a quantitative approach to historical language choice, of course, cannot provide solid answers to all questions. In fact, not all social or sociolinguistic factors potentially conditioning the choice of language were equally suitable to be incorporated into our methodological framework. The additional sociolinguistic variable of age, for instance, is more difficult to integrate in a balanced manner, especially when following the selection protocol presented in Section 3. However, in order to shed more light on the intriguing differences in correspondence between children and parents, the age of children may be a revealing factor. It should be noted, though, that this requires the availability of letters distributed across writers' lifespans (which is not often the case). Similar challenges apply to the variables of social class and education, which are more or less stable in our dataset, but may be worth exploring in future research. Further factors not covered and less suitable for a quantitative approach include individual attitudes and personal states of mind (Ruberg 2011: 69–70), political factors (van Strien-Chardonneau 2018: 77–81) or, in fact, more functional constraints like the level of proficiency of senders and addressees (Head 1995: 591–592).

Ultimately, qualitative and quantitative methods to a complex phenomenon like language choice are never mutually exclusive but best seen as complementary approaches (cf. Nevalainen & Raumolin-Brunberg 2017: 244). However, the quantitative results from our analyses, for instance on regional differences, gender and constellations of family members, allow us to interpret previous (and future) qualitative findings against the background of reliable baseline data on Dutch-French language choice in the specific domain of private life.

Funding sources
The research was supported by the Dutch Research Council (NWO), project Pardon my French? Dutch-French Language Contact in The Netherlands, 1500–1900.

REFERENCES

Argent, G., Rjéoutski, V. & Offord, D. (2014): European francophonie and a framework for its study. In: V. Rjéoutski, G. Argent & D. Offord (eds.), *European Francophonie: The Social, Political and Cultural History of an International Prestige Language*. Berlin, Bern, Bruxelles, New York, Oxford, Warszawa, Wien: Peter Lang, 1–32.

Assendelft, B., Rutten, G. & van der Wal, M. (forthcoming): Tracing Frenchification: A sociolinguistic analysis of French loan suffixes in a historical corpus of Dutch. In: R. Franceschini, M. Hüning & P. Maitz (eds.), *Historische Mehrsprachigkeit: Europäische Perspektiven*. Berlin: De Gruyter.

Böhm, M. (2010): *Sprachenwechsel: Akkulturation und Mehrsprachigkeit der Brandenburger Hugenotten vom 17. bis 19. Jahrhundert*. Berlin, New York: De Gruyter.

de Meyer, G. M. (1974): Latijn en volkstaal, schrijftaal en spreektaal in niet-literaire teksten [Latin and vernacular, written and spoken language in non-literary texts]. *Bijdragen en Mededelingen betreffende de Geschiedenis der Nederlanden* 89, 1–15.

de Vooys, C. (1970): *Geschiedenis van de Nederlandse taal* [History of the Dutch Language]. Groningen: Wolters-Noordhoff.

Dodde, N. (2020): *Franse scholen in Nederland: Ontstaan en ontwikkeling vanaf de vijftiende eeuw tot het midden van de negentiende eeuw* [French Schools in the Netherlands: Origin and Development from the Fifteenth Century to the Mid-Nineteenth Century]. Oud-Turnhout, 's-Hertogenbosch: Gompel & Svacina.

Elspaß, S. (2012): The use of private letters and diaries in sociolinguistic investigation. In: J. M. Hernández-Campoy & J. C. Conde-Silvestre (eds.), *The Handbook of Historical Sociolinguistics*. Chichester: Blackwell, 156–169.

Fishman, J. A. (1965): Who speaks what language to whom and when? *La Linguistique* 1, 67–88.

Frijhoff, W. (1989): Verfransing? Franse taal en Nederlandse cultuur tot in de revolutietijd [Frenchification? French language and Dutch culture until the Revolutionary period]. *Bijdragen en Mededelingen betreffende de Geschiedenis der Nederlanden* 104, 592–609.

Frijhoff, W. (2003): Uncertain brotherhood: The Huguenots in the Dutch Republic. In: B. Van Ruymbeke & R. Sparks (eds.), *Memory and Identity: The Huguenots in France and the Atlantic Diaspora*. Columbia: University of South Carolina Press, 128–171.

Frijhoff, W. (2015): Multilingualism and the challenge of Frenchification in the early modern Dutch Republic. In: G. Rutten & R. Vosters (eds.), *Past, Present and Future of a Language Border: Germanic-Romance Encounters in the Low Countries*. Berlin, München, Boston: De Gruyter Mouton, 115–140.

Gardner-Chloros, P. (2018): Historical and modern studies of codeswitching: A tale of mutual enrichment. In: P. Pahta, J. Skaffari & L. Wright (eds.), *Multilingual Practices in Language History: English and Beyond*. Berlin, Boston: De Gruyter Mouton, 19–36.

Gretchanaia, E., A. Stroev & Viollet, C. (eds.) (2012): *La francophonie européenne aux XVIIIe–XIXe siècles*. Berlin, Bern, Bruxelles, New York, Oxford, Warszawa, Wien: Peter Lang.

Head, R. C. (1995): A plurilingual family in the sixteenth century: Language use and linguistic consciousness in the Salis family correspondence, 1580–1610. *The Sixteenth Century Journal* 26, 577–593.

Joby, C. (2014): *The Multilingualism of Constantijn Huygens (1596–1687)*. Amsterdam: Amsterdam University Press.

Kessels-van der Heijde, M. (2002): *Maastricht, Maestricht, Mestreech: De taalverhoudingen tussen Nederlands, Frans en Maastrichts in de negentiende eeuw* [Maastricht, Maestricht, Mestreech: Language Relations between Dutch, French and Maastricht in the Nineteenth Century]. Hilversum: Verloren.

Kessels-van der Heijde, M. (2015): The use of languages in Maastricht in the nineteenth century: the press and family archives. In: C. Peersman, G. Rutten & R. Vosters (eds.), *Past, Present and Future of a Language Border: Germanic-Romance Encounters in the Low Countries*. Berlin, München, Boston: De Gruyter Mouton, 171–204.

Kloek, J. & Mijnhardt, W. (2001): *1800: Blauwdrukken voor een samenleving* [1800: Blueprints for a National Community]. The Hague: Sdu.

Knuttel, W. P. C. (1978): *Catalogus van de pamfletten-verzameling berustende in de Koninklijke Bibliotheek, 1486–1853* [Catalogue of the pamphlet collection kept in the Royal Library of the Netherlands, 1486–1853]. Reprint with corrections and an introduction by H. van der Hoeven. Originally printed in 1889–1920. Utrecht: HES Publishers.

Krogull, A., Puttaert, J. & Rutten, G. (submitted): Verfransing in de achttiende-eeuwse Noordelijke Nederlanden? Empirisch onderzoek naar taalkeuze in egodocumenten en pamfletten [Frenchification in the eighteenth-century Northern Netherlands? Empirical research into language choice in ego-documents and pamphlets].

Leemans, I. & Johannes, G.-J. (2013): *Worm en donder: Geschiedenis van de Nederlandse literatuur 1700–1800: de Republiek* [Worm and thunder: History of Dutch literature 1700–1800: the Republic]. Amsterdam: Bert Bakker.

Lucassen, L. & de Vries, B. (2001): The rise and fall of a West European textile-worker migration system: Leiden, 1586–1700. In: G. Gayot & P. Minard (eds.), *Les ouvriers qualifiés de l'industrie (XVIe–XXe siècle): Formation, emploi, migrations*. Lille: Université Charles de Gaulle, 23–42.

Myers-Scotton, C. (1993): *Duelling Languages: Grammatical Structures in Code-Switching*. Oxford: Oxford University Press.

Nevalainen, T. & Raumolin-Brunberg, H. (2017): *Historical Sociolinguistics: Language Change in Tudor and Stuart England*. Abingdon, New York: Routledge.

Noordam, D. J. (2003): Demografische ontwikkelingen [Demographic developments]. In: R. van Maanen & S. Groeneveld (eds.), *Leiden: De geschiedenis van een Hollandse stad, Deel 2: 1574–1795* [Leiden: The History of a Dutch City, Volume 2: 1574–1795]. Leiden: Stichting Geschiedschrijving Leiden, 42–53.

Offord, D. (2020): Sociolinguistics and history: An interdisciplinary view of bilingualism in Imperial Russia. *Journal of Historical Sociolinguistics* 6, 20180019.

Offord, D., Rjéoutski, V. & Argent, G. (eds.) (2018): *The French Language in Russia: A Social, Political and Cultural History of an International Prestige Language*. Berlin, Bern, Bruxelles, New York, Oxford, Warszawa, Wien: Peter Lang.

Pahta, P., Skaffari, J. & Wright, L. (eds.) (2018): *Multilingual Practices in Language History: English and Beyond*. Berlin, Boston: De Gruyter Mouton.

Peersman, C. (2012): Written vernaculars in Medieval and Renaissance times. In: J. M. Hernández-Campoy & J. C. Conde-Silvestre (eds.), *The Handbook of Historical Sociolinguistics*. Chichester: Blackwell, 639–654.

Peersman, C., Rutten, G. & Vosters, R. (eds.) (2015): *Past, Present and Future of a Language Border: Germanic-Romance Encounters in the Low Countries*. Berlin, München, Boston: De Gruyter Mouton.

Prevenier, W. & de Hemptinne, T. (2005): La Flandre au moyen âge: Un pays de trilinguisme administratif. In: O. Guyotjeannin (ed.), *La langue des actes*. Paris: ELEC, éditions en ligne de l'Ecole des Chartes, 1–16.

Rjéoutski, V., Argent, G. & Offord, D. (eds.) (2014): *European Francophonie: The Social, Political and Cultural History of an International Prestige Language*. Berlin, Bern, Bruxelles, New York, Oxford, Warszawa, Wien: Peter Lang.

Rjéoutski, V. & Frijhoff, W. (2018): *Language Choice in Enlightenment Europe. Education, Sociability, and Governance*. Amsterdam: Amsterdam University Press.

Ruberg, W. (2011): *Conventional Correspondence: Epistolary Culture of the Dutch Elite, 1770–1850*. Leiden: Brill.

Rutten, G., Vosters, R. & van der Wal, M. (2015): Frenchification in discourse and practice: Loan morphology in Dutch private letters of the eighteenth and nineteenth centuries. In: C. Peersman, G. Rutten & R. Vosters (eds.), *Past, Present and Future of a Language Border: Germanic-Romance Encounters in the Low Countries*. Berlin, München, Boston: De Gruyter Mouton, 143–170.

Rutten, G. & van der Wal, M. (2014): *Letters as Loot: A Sociolinguistic Approach to Seventeenth- and Eighteenth-Century Dutch*. Amsterdam, Philadelphia: John Benjamins.

Schendl, H. (2012): Multilingualism, code-switching, and language contact in historical sociolinguistics. In: J. M. Hernández-Campoy & J. C. Conde-Silvestre (eds.), *The Handbook of Historical Sociolinguistics*. Chichester: Blackwell, 520–533.

Sleiderink, R. (2010): From Francophile to Francophobe: The changing attitude of Medieval Dutch authors towards French literature. In: C. Kleinhenz & K. Busby (eds.), *Medieval Multilingualism: The Francophone World and its Neighbours*. Turnhout: Brepols, 127–143.

Stell, G. (2019): Code-switching. In: J. Darquennes, J. C. Salmons & W. Vandenbussche (eds.), *Language Contact: An International Handbook, Volume 1*. Berlin, Boston: De Gruyter Mouton, 159–172.

Thomas, J. (2017): *"Vous êtes hombre de bien": A Study of Bilingual Family Letters to and from Colonial Louisiana, 1748–1867*. PhD dissertation. Berkeley: University of California.

van den Toorn, M., Pijnenburg, W., Leuvensteijn, A. van & van der Horst, J. (eds.) (1997): *Geschiedenis van de Nederlandse taal* [History of the Dutch language]. Amsterdam: Amsterdam University Press.

van der Horst, J. (2008): *Geschiedenis van de Nederlandse syntaxis* [History of Dutch Syntax]. Leuven: Universitaire Pers Leuven.

van der Sijs, N. (2002): *Chronologisch woordenboek: De ouderdom en herkomst van onze woorden en betekenissen* [Chronological Dictionary: The Age and Origin of Our Words and Meanings]. Amsterdam, Antwerpen: Veen.

van der Wal, M. (1994): De opstand en de taal: Nationaal bewustzijn en het gebruik van het Nederlands in het politieke krachtenveld [The revolt and the language: National consciousness and the use of Dutch in the political arena]. *De Zeventiende Eeuw* 10, 110–115.

van der Wal, M. (2018): Early Modern migrants in a language contact setting: characteristics of the Dutch Heusch correspondence (1664–1665). *Journal of Historical Sociolinguistics* 4, 253–280.

van der Wal, M. & van Bree, C. (2014): *Geschiedenis van het Nederlands* [History of Dutch]. Houten: Spectrum.

van Strien-Chardonneau, M. (2014): The use of French among the Dutch elites in eighteenth-century Holland. In: V. Rjéoutski, G. Argent & D. Offord (eds.), *European Francophonie: The Social, Political and Cultural History of an International Prestige Language*. Berlin, Bern, Bruxelles, New York, Oxford, Warszawa, Wien: Peter Lang, 145–173.

van Strien-Chardonneau, M. (2018): Practice and functions of French as a second language in a Dutch patrician family: The van Hogendorp family (Eighteenth–early nineteenth centuries). In: V. Rjéoutski & W. Frijhoff (eds.), *Language Choice in Enlightenment Europe: Education, Sociability, and Governance*. Amsterdam: Amsterdam University Press, 65–85.

van Strien-Chardonneau, M. & Kok Escalle, M. C. (eds.) (2017): *French as Language of Intimacy in the Modern Age*. Amsterdam: Amsterdam University Press.

Verschaffel, T. (2017): *De weg naar het binnenland: Geschiedenis van de Nederlandse literatuur 1700–1800: de Zuidelijke Nederlanden* [The Road to the Inland: History of Dutch literature 1700–1800: the Southern Netherlands]. Amsterdam: Bert Bakker.

Vogl, U. (2015): Standard language ideology and the history of Romance-Germanic encounters. In: C. Peersman, G. Rutten & R. Vosters (eds.), *Past, Present and Future of a Language Border: Germanic-Romance Encounters in the Low Countries*. Berlin, München, Boston: De Gruyter Mouton, 61–87.

Willemyns, R. (1994): Taalpolitiek in de Bourgondische tijd [Language politics in the Burgundian era]. *Verslagen en Mededelingen van de Koninklijke Academie voor Nederlandse Taal – en Letterkunde* 104, 162–177.

Wright, S. (2016): *Language Policy and Language Planning: From Nationalism to Globalization*. 2nd edition. Basingstoke: Palgrave Macmillan.

ZUSAMMENFASSUNG

Auf dem Weg zu einem methodologischen Rahmen für die historische Sprachenwahl: Niederländisch und Französisch in den Niederlanden (1800–1899)

Das Ziel des vorliegenden Beitrags ist es, den Forschungsgegenstand historische Sprachenwahl aus einer quantitativen Perspektive zu betrachten. Dabei soll gezeigt werden, dass eine solide Grundlage empirischer Forschungsergebnisse, basierend auf umfangreichen Datensätzen, eine dringend benötigte Ergänzung zu den bisherigen, größtenteils qualitativen Beobachtungen darstellt. Wir stellen einen methodologischen Rahmen vor, der die Untersuchung soziolinguistischer Faktoren, die die Sprachenwahl im privaten Bereich möglicherweise beeinflussen, ermöglichen soll. Um das Potenzial unserer Methodologie zu veranschaulichen, präsentieren wir eine Fallstudie zur niederländisch-französischen Sprachenwahl in den Niederlanden des 19. Jahrhunderts, mit dem Fokus auf privater Familienkorrespondenz. Unser Beitrag zeigt, dass ein sorgfältiges Selektionsverfahren bei umfangreichen historischen Datensätzen ausschlaggebend ist um eine ausgewogene Repräsentation von Sprachenwahl zu erreichen. Im Hinblick auf unsere empirischen Analysen erweist sich die Rolle des Französischen in Privatbriefen als relativ gering, während das Niederländische deutlich überwiegt. Dennoch werden interessante Muster sichtbar, wenn man regionale Unterschiede, Geschlechterkonstellationen und Familienbeziehungen als Faktoren in Betracht zieht. Die quantitativen Forschungsergebnisse dieser Studie bieten somit auch einen interpretativen Rahmen für qualitative Studien zur historischen Sprachenwahl im niederländisch-französischen Kontext und darüber hinaus.

Leiden University Centre for Linguistics
Postbus 9515, 2300 RA Leiden, Netherlands
<j.puttaert@hum.leidenuniv.nl>
<a.krogull@hum.leidenuniv.nl>
<g.j.rutten@hum.leidenuniv.nl>