Policy versus Practice. Language variation and change in eighteenth- and nineteenth-century Dutch

Krogull, A.

CHAPTER 3

Theoretical framework

1 The field of historical sociolinguistics

The research presented in this dissertation is embedded in the theoretical framework of historical sociolinguistics. Over the past three decades, this field of study has developed into a mature and well-established linguistic (sub)discipline, which focuses on the reconstruction and study of language in its historical and social context. Historical sociolinguistics is described as "a hybrid subfield subsisting on the interdisciplinary character of sociolinguistic methodology" (Conde-Silvestre & Hernández-Campoy 2012: 1), drawing on findings and principles of (modern-day) sociolinguistics and historical linguistics, as well as social history. Nevalainen & Raumolin-Brunberg (2003: 10) point out that

> [r]econstructing how language changes diffuse socially is one of the major tasks, if not *the* major task, of historical sociolinguists. This is an area where the 'historical' in 'historical sociolinguistics' is connected not only with historical linguistics but also with social history.

From an advanced cross-disciplinary perspective, historical sociolinguistics is also connected with even more associated fields of study, such as corpus linguistics, philology and dialectology (Nevalainen & Raumolin-Brunberg 2012: 27).

Suzanne Romaine's (1982) *Socio-historical Linguistics: Its status and methodology*, a systematic study of relative markers in Middle Scots, is usually considered as the initiation of historical sociolinguistics as a specialised field of study (Auer et al. 2015: 2). Terminologically, her book was the first to include *sociohistorical linguistics* in its title. The same term was also used by Tieken-Boon van Ostade (1987) in her study of auxiliary *do* in eighteenth-century English. The alternative term *historical sociolinguistics*, which has been established as the most common term internationally, first appeared in Milroy's (1992) *Linguistic Variation and Change: on the Historical Sociolinguistics of English*. It also occurred in the title of a theoretical chapter on historical sociolinguistics by Nevalainen (1996). These pioneering works were largely inspired by the empirical research methods in modern-day sociolinguistics and addressed the need to apply such an approach to language history (Rutten et al. 2014b: 1).

It should be noted that there had been many publications on what would nowadays be considered as historical-sociolinguistic research before Romaine's (1982) study, without using either of the terms *sociohistorical linguistics* or *historical sociolinguistics* (Auer et al. 2015: 2). In a handbook chapter on historical sociolinguistics, Romaine (2005: 1696) explicitly refers to the influential paper by Weinreich et al. (1968), which already emphasised the need to incorporate external

factors into language change theory in the late 1960s and thus "laid the foundations for an approach to language that was inherently historical and social" (Auer et al. 2015: 2). While the above-mentioned seminal works in historical sociolinguistics suggest a strong focus of the field on the language history of English, the 1980s and 1990s also saw the parallel development of a German tradition (e.g. Mattheier 1988; Mihm 1998).

The first decade of the twenty-first century witnessed the gradual development and expansion of historical sociolinguistics. Nevalainen & Raumolin-Brunberg's (2003) book on language change in Tudor and Stuart England can be regarded as an introduction to the field of historical sociolinguistics, both explaining sociolinguistic concepts and demonstrating methods to apply these concepts to language in the past. Furthermore, two handbook chapters (Romaine 2005; Roberge 2006) as well as the paper by Willemyns & Vandenbussche (2006) manifested historical sociolinguistics as a thriving field of study. The same period also saw the establishment of the *Historical Sociolinguistic Network* (HiSoN), which has organised numerous international conferences, workshops and annual summer schools ever since.

A landmark for the field was undoubtedly the publication of *The Handbook of Historical Sociolinguistics* (Hernández-Campoy & Conde-Silvestre 2012), which testified to the multi-faceted research on different topics, language areas and time periods. The institutionalisation of historical sociolinguistics as an independent discipline within linguistics was further advanced by the launch of two academic book series, viz. *Advances in Historical Sociolinguistics* (John Benjamins), edited by Marijke van der Wal and Terttu Nevalainen, and *Historical Sociolinguistics. Studies on Language and Society in the Past* (Peter Lang), edited by Nils Langer, Stephan Elspaß, Joseph Salmons and Wim Vandenbussche. In 2015, publisher De Gruyter launched the *Journal of Historical Sociolinguistics*, a new academic journal specifically devoted to historical-sociolinguistic research, which is edited by Gijsbert Rutten, Anita Auer, José del Valle, Rik Vosters and Simon Pickl.

In the following, Section 2 briefly outlines the basic principles and challenges in historical sociolinguistics. The different approaches to language histories, both from a traditional perspective and from below, will be discussed in Section 3. Section 4 provides an overview of historical-sociolinguistic data collections and corpora. The interrelated topics of prescriptivism, language planning and policy (and their possible effects on actual language usage) will be addressed in Section 5.

2 Principles and challenges

As mentioned in Section 1, historical sociolinguistics largely draws on insights and concepts from present-day sociolinguistics, assuming that the fundamental principles and mechanisms of language variation and change do not change over time. This assumption is based on the so-called *uniformitarian principle*, which, according to Labov (1972: 275), implies that "the forces operating to produce

linguistic change today are of the same kind and order to magnitude as those which operated five or ten thousand years ago". Although Bergs (2012: 96) agrees on the fact "that language must always have been variable, that different social groups and genders had different ways of speaking, and that people have always been aware of these differences", he also warns of what he refers to as *ideational anachronism*, i.e. the danger of hastily transposing modern concepts like social class or gender to historical settings. Therefore, Auer et al. (2015: 5) argue that

> it is the task of historical sociolinguists to reconstruct a broad picture of the social context in which the language varieties under investigation were used, drawing on the inductive method to identify the social conditions of language variation and change, ensuring empirical, social and historical validity.

While modern sociolinguistics is primarily based on spoken language data, language use in earlier periods – at least before the advent of audio recordings – can only be accessed by written sources. Addressing the limitations of historical (written) data, which are naturally fragmented records of the past and often preserved by chance, Labov (1994: 11) remarks that historical linguistics can be described as "the art of making the best use of bad data". In that sense, the well-known *bad data problem* might be even more challenging for historical *socio*linguists, as background information on the informants' age, gender, regional origin and mobility, education, and so on, are often hard to find and reconstruct. Nevertheless, the dependence on supposedly 'bad' data [14] should not be overstressed and can, at least to a certain extent, be compensated by "systematicity in data collection, extensive background reading and good philological work", as Nevalainen & Raumolin-Brunberg (2003: 26) point out. They further admit that these tasks are "demanding and time-consuming, but by no means unrealizable" (ibid.).

In addition to their fragmented nature, written sources from the past share the fundamental bias that there are no direct records of the spoken language in historical writing. However, it is often claimed that speech is primary and thus essential for the understanding of language variation and change processes (Auer et al. 2015: 7; cf. also Schneider 2013: 57). This means that for a thorough study of variation and change that has to rely on written documents only, the relationship between speech and writing has to be critically (re)assessed. Historical sociolinguists in particular have to tackle the questions "how oral registers can be reconstructed from written sources, and thus […] how written and oral language are interrelated" (Auer et al. 2015: 7).

[14] Nevalainen & Raumolin-Brunberg (2003: 27) further remark that "[t]he comments on bad data easily lead to an impression that historical texts would be inferior to present-day material in every respect". They argue that this is "not necessarily the case", listing a number of advantages of these texts. In contrast to modern sociolinguistics, historical data have the major advantage of real-time analyses of language change. Moreover, these texts represent genuine communication from the past, unaffected by the participation of the researcher.

In search of suitable, i.e. 'oral'-like written data for investigating language variation and change in the past, the field of historical sociolinguistics heavily relies on the influential frameworks by Biber (1988) and Koch & Oesterreicher (1985), who criticise the rigid distinction between speech and writing. The traditional dichotomy of spoken and written language, as Elspaß (2012: 157) remarks, is too "simplistic and even misleading". Suggesting an alternative and more sophisticated model, Koch & Oesterreicher (1985) therefore differentiate between the *medium* and the *conception* of language. On the one hand, the linguistic medium is a strict dichotomy between the phonic and the graphic code. On the other hand, the conception of language distinguishes linguistic utterances on the basis of the communicative strategies they utilise. In contrast to the medial dichotomy, the conception has to be regarded as a continuum, on which all types of texts and genres can be located according to their specific situational and communicative parameters. The two poles on this conceptual continuum are referred to as the 'language of immediacy' (*Sprache der Nähe*) and the 'language of distance' (*Sprache der Distanz*). The pole of immediacy represents texts that are prototypically 'oral' (dialogic, informal, unplanned, and so on, like an intimate conversation), whereas texts close to the pole of distance are prototypically 'literate' (monologic, formal, planned, and so on, like a legal contract) (Elspaß 2012: 157). In fact, some (medially) spoken genres, such as sermons, mainly utilise literate strategies, whereas some (medially) written genres, such as private letters, are relatively close to the 'oral' pole of the continuum (Koch & Oesterreicher 1985: 23). Elspaß (2012: 157) highlights the importance of the linguistic conception as opposed to the linguistic medium:

> The distinction between the conceptual poles of the continuum is primary to an understanding of the 'orality' or 'writtenness' of language. The medium is secondary, as a written text can be read out and a recording of a spoken text can be transcribed at any time.

For historical sociolinguists, who seek to find written data that reflect the spoken language, or rather orality, as much as possible, genres close to the 'language of immediacy' are thus highly valuable sources. Private letters are generally characterised as one of the most oral written genres (cf. Biber 1988), which make them "first-class primary data" for historical-sociolinguistic research (Nevalainen & Raumolin-Brunberg 2003: 29). More generally, so-called ego-documents, a cover term for texts from the private sphere "in which an author writes about his or her own acts, thoughts and feelings" (Dekker 2002: 14), for instance letters, diaries and travelogues, have proven to be useful sources as they are "usually close to speech and relatively unaffected by conventions of writing" (Auer et al. 2015: 7; cf. also Chapter 4). In historical sociolinguistics, these comparatively 'oral'-like ego-documents in particular are at the heart of the so-called *language history from below*, an alternative approach to traditional language histories, which will be outlined in Section 3.

3 Language histories from different perspectives

3.1 Traditional language histories

For most European languages, including Dutch, traditional language historiography was first and foremost concerned with the unification and standardisation process of a given language. Until the late twentieth century, language histories were strongly dominated by the teleological view on language change as "the inexorable march towards a uniform standard" (Elspaß 2007: 3), presenting a more or less linear development from medieval variation to present-day uniformity, largely restricted to the evolution of an ideal and invariant standard variety.

Language histories are, in fact, greatly influenced by an underlying standard language ideology, defined by Lippi-Green (2012: 67) as "a bias toward an abstract, idealized homogeneous language, which is imposed and maintained by dominant bloc institutions and which names as its model the written language". The origins of this ideology go back to the eighteenth century as a side effect of the period of nation building (cf. Chapter 2; cf. also Rutten 2016b). Referring to the standard language ideology, Milroy (2012: 579) emphasises that its influence "on traditional historical descriptions of major modern languages cannot be overstated". The consequences of such a single-stranded approach are particularly noticeable in introductory textbooks on the subject, which easily give the impression that a language's post-medieval history is equivalent to the history of its standard variety. Similarly, Watts (2012: 585) argues that traditional language historiography is grounded on "an implied teleology […] that standard languages are the only valid objects of study for a language history". He metaphorically calls this approach a tunnel vision view, "which projects from a source domain of the restriction, unidirectionality, and perhaps even darkness of passage through a tunnel onto the target domain of the abstract concept of language history" (Watts 2012: 585; cf. also Watts & Trudgill 2002). By narrowing historical descriptions of languages down to their standardisation process, that is, ignoring or rejecting the existence of non-standard(ised) forms and varieties (also referred to as erasure), incomplete language histories are written (Watts 2012: 579).

In historical sociolinguistics, the teleological view on language history has often been criticised for being "one-sided, partial, biased, largely based on a limited collection of text sources" (Rutten et al. 2014b: 2). Indeed, language historiography primarily described the evolution of printed languages (Elspaß 2012: 156). Partly motivated by the underlying standard language ideology, partly due to the wider availability and easier accessibility of (printed) data, language historiographers mainly focused on literary and formal texts from the higher registers (van der Wal 2006). However, as Schneider (2013: 59) remarks, these texts "normally display categorical, invariant usage and fail to reflect natural speech behaviour and associated processes". Furthermore, literary and formal writings have a strong social and gender bias, as they were almost exclusively produced by well-educated men from the upper ranks of society. According to Elspaß (2007: 4-5), this restricted group of elite writers probably constituted no more than five per cent of

the population, at least in the case of nineteenth-century Germany. In other words, a substantial part of language use and language users of the past is not represented in these conventional language histories.

Examples of this traditional perspective on the history of Dutch, with a strong focus on standardisation and the increasing uniformity in literary and formal texts, can be found in van der Sijs (2004) and van der Wal & van Bree (2008), among many others. The specific case of negation in Dutch shows that these linguistic histories often portrayed a clear development in which single negation, as opposed to bipartite negation, had been selected as the norm by the mid-seventeenth century and then became standardised (e.g. van der Sijs 2004: 534-537; van der Wal & van Bree 2008: 217-218). This might indeed be the case in more formal registers and texts produced by literary authors. However, new historical-sociolinguistic insights gained from seventeenth-century private letters give evidence that "[i]n the spoken language, or perhaps more generally in informal registers, both spoken and written, bipartite negation remained in use" (Rutten 2016c: 204; cf. also Rutten & van der Wal 2014: ch. 10). This concrete counterexample illustrates that the study of a limited collection of texts can draw an incomplete picture of the linguistic past. As Elspaß (2007: 4) argues, the "neglect of texts 'below' the surface of printed language […] has led to a language historiography in which a major part of both the language community (i.e. those writers with no access to printing) and their written language production is simply not represented".

3.2 Language histories from below

In order to provide an alternative approach to language history, historical sociolinguists have suggested the *language history from below* (e.g. Elspaß et al. 2007). To begin with, it should be emphasised that the term *from below* does not correspond with the Labovian concepts indicating linguistic phenomena from above or below the level of awareness (Labov 1994: 78). Furthermore, *from below* as a specific perspective on language history does not refer to linguistic changes starting in the lower social classes (cf. also Rutten & van der Wal 2014: 5). In line with previous historical-sociolinguistic studies, and following Elspaß's (2007: 5) definition, I will use the term *from below* in reference to "the social ranks below the highest social class and to texts representing everyday language that could thus be considered as below formal registers such as the language of literature".

The central aim of the *language history from below* approach is to fill in the substantial gaps – or *witte vlekken* 'blank areas', as van der Wal (2006) puts it – in traditional language historiography, and thus to contribute to a more complete reconstruction of languages in the past. Auer et al. (2015: 6) point out that "[o]ne of the core concerns of historical sociolinguistics […] is the effort to overcome the social bias connected to class, education and literacy inherent in written sources that has afflicted historiography".

With respect to textual sources, the change of perspective involves a shift from printed, often literary and formal texts to handwritten, more 'oral'-like documents. Although we do not have direct records of historical spoken language, as noted in Section 2 before, there are certain genres which reflect the 'language of immediacy' as much as possible and can, therefore, be considered as valuable sources for historical-sociolinguistic research. The approach *from below* usually takes advantage of these ego-documents, comprising various types of texts that have been "important in people's private lives and personal experiences" (Rutten et al. 2014b: 1). It has repeatedly been argued in historical-sociolinguistic research (e.g. Nevalainen & Raumolin-Brunberg 2003, Elspaß 2005, Rutten & van der Wal 2014) that in order to gain access to authentic language use in historical contexts, we need to find those textual sources that are "as close to actual speech as possible, only in written form" (Sevic 1999: 34). In Koch & Oesterreicher's (1985) terms, ego-documents like private letters, diaries and travelogues are considered to be conceptually more oral and thus closer to the 'language of immediacy' (Section 2) than the conceptually much more 'literate' writings, on which conventional language histories were grounded.

Moreover, ego-documents also have the advantage that they were produced by much wider parts of the language community than just the elite. Private letters and other types of ego-documents were written by both men and women from different social classes (including the lower and middle classes), from different regions, and so on. Forming the basis of the *from below* approach, ego-documents thus offer a remarkable opportunity of filling in the gaps left by language historians. Elspaß (2012: 161) argues that this approach is "a necessary counterweight" to the teleological historiography dominated by the standardisation perspective.

Nonetheless, it has also been remarked more recently by Rutten et al. (2014b: 2; cf. also Fairman 2007) that

> purposely and explicitly leaving aside the more 'standard'-like textual sources found in print, in literature, in elite documents, and setting aside the possible influence of supraregional writing conventions, language norms and prescriptions, may run the risk of presenting another one-sided view of language history.

Thus, in order to avoid a description of language history that becomes just as one-sided and biased as the traditional approach, historical sociolinguists have called for an integrated approach to language history, combining the two perspectives from above (i.e. in the traditional sense) *and* below (Rutten et al. 2014b: 2):

> Based on the considerable research tradition in historical sociolinguistics that has come into existence over the past few decades, the time has now come to integrate both perspectives, and to reassess the importance of language norms, standardization and prescription on the basis of sound empirical studies of large corpora of texts.

In this dissertation, such an integrated perspective is taken in order to investigate the effectiveness of the national language policy and officially standardised language norms on actual usage in the Northern Netherlands. In other words, central issues typically known from traditional language histories, such as standardisation, language norms and prescriptions, will be reassessed on the basis of new insights gained from the perspective from below. In concrete terms, the possible 'success' of the Dutch language policy in the early 1800s, with its officialised regulations for spelling and grammar, will be examined on the basis of a newly compiled corpus of eighteenth- and nineteenth-century Dutch that makes use of handwritten ego-documents from the private sphere (in line with recent historical-sociolinguistic research), but also integrates more 'standard'-like printed texts (cf. also Elspaß & Niehaus 2014). In this respect, I consider the different perspectives on language history discussed in this section as complementary rather than contradictory. The corpus and methodology applied in this dissertation will be presented in more detail in Chapter 4.

4 Data and corpora

Most studies in the field of historical sociolinguistics are based on empirical data, usually combining quantitative and qualitative findings. In fact, the collection of suitable data has played a central role in historical-sociolinguistic research ever since the advent of corpus linguistics in the 1980s. A number of corpora have been compiled for various languages and time periods.

In the early days, the focus was on diachronic multi-genre corpora, which allowed a systematic comparison of different registers and genres over time. Without any doubt, genre was considered as "the key external variable in sociolinguistically informed studies of language change" (Nevalainen & Raumolin-Brunberg 2012: 23). An example of such a multi-genre corpus is the *Helsinki Corpus of English Texts*, commonly regarded as the first historical corpus of English (cf. <http://www.helsinki.fi/varieng/CoRD/corpora/HelsinkiCorpus>). The corpus counts more than 1.5 million words in total, comprising text samples from the Old, Middle and Early Modern English periods, which are enriched by metadata on the selected texts and their authors.

Another diachronic multi-genre corpus of English is *ARCHER*: *A Representative Corpus of Historical English Registers*, initiated by Douglas Biber and Edgar Finegan in the 1990s. In its most recent version (*ARCHER* 3.2, completed in 2013, cf. <http://www.projects.alc.manchester.ac.uk/archer>), the corpus comprises 3.3 million words of British and American English texts from the seventeenth to the twentieth centuries, covering a range of both prototypically written texts, for instance medical and scientific prose, but also more speech-like genres like letters and diaries.

Following the model of historical corpora of English, especially *ARCHER*, Martin Durrell's *GerManC* project (Durrell et al. 2012) aims to provide a representative sample of German in the period between 1650–1800. The *GerManC*

corpus comprises 800,000 words, covering texts from eight genres, including more orally oriented genres like personal letters, sermons and drama, as well as more print-oriented genres like scientific and legal texts. Adding a spatial dimension to the corpus, the selected texts derive from five different regions in the language area (cf. <http://www.llc.manchester.ac.uk/research/projects/germanc>).

Particularly from the 2000s onwards, historical sociolinguistics has witnessed a shift from textually balanced multi-genre corpora towards single-genre corpora (Nevalainen & Raumolin-Brunberg 2003: 27). With respect to the remarkable value of ego-documents for historical sociolinguistics[15], as discussed in Section 3, it is "no coincidence that many of the corpora explored for this research consist of private letters" (Rutten & van der Wal 2014: 4; cf. also Elspaß 2005, 2012: 47-48). One of the pioneering historical corpora of letter data is the *Corpus of Early English Correspondence* (CEEC) (Nevalainen & Raumolin-Brunberg 2003), compiled at the University of Helsinki under the direction of Terttu Nevalainen. The central aim of the project was to test how the methods used in present-day sociolinguistics could be applied to historical data. The original 1998 version of the CEEC contains c. 2.6 million words, comprising around 6,000 letters by 778 informants, written in the period 1410–1681. More recently, the corpus was extended to a time span of four centuries (1400-1800), and to 11,819 letters by 1,066 writers (Palander-Collin et al. 2009: 14). For the field of historical sociolinguistics, the CEEC can certainly be considered ground-breaking as it was the first corpus which was tagged with fairly elaborate sociolinguistic metadata on the informants (cf. <http://www.helsinki.fi/varieng/CoRD/corpora/CEEC>).

For the history of German, Stephan Elspaß's (2005) corpus of nineteenth-century emigrant letters, written by men and women from the lower and lower-middle ranks of society, testifies to the relevance of the approach from below. The corpus of emigrant letters (*Auswandererbriefe*) comprises around 700,000 words, 648 letters and 273 writers (Elspaß 2012: 47). Also integrating a geographical dimension, the corpus represents data from different regions of the investigated language area.

In order to fill in the major gaps in the language history of Dutch, the *Letters as Loot* corpus (Rutten & van der Wal 2014: 16; cf. also Nobels 2013, Simons 2013) was compiled at Leiden University under the direction of Marijke van der Wal. It contains private letters from the late seventeenth and late eighteenth centuries (i.e. two diachronic cross-sections: 1660s–1670s, 1770s–1780s), which were confiscated as so-called 'Prize Papers' during the Anglo-Dutch Wars and have been kept in the British National Archives. The *Letters as Loot* corpus counts 424,000 words and a selection of 933 private letters produced by 716 writers. These letters are a unique collection, as they were produced by both men and women from all ranks of society and different regional backgrounds. Furthermore, the electronic *Letter as Loot* corpus is lemmatised, tagged for parts of speech, and includes detailed metadata (cf. <http://brievenalsbuit.inl.nl>).

[15] See Elspaß (2012: 162-163) for an overview of various corpora of ego-documents.

When we look beyond the Germanic languages, France Martineau's *Corpus de français familier ancient* is certainly worth mentioning, as it is "a unique sociohistorical French megacorpus of 20,000 non-literary texts, such as family letters, diaries and account books" (Martineau 2013: 132). The corpus is based on original documents found in archives in France and North America, dating from the seventeenth to the early twentieth century. For Portuguese, Rita Marquilhas's *CARDS – Unknown Letters* project compiled a historical digital archive of Portuguese letters, spanning four centuries from the sixteenth to the nineteenth centuries. The succeeding project *P.S. Post Scriptum* focuses on private letters from different social backgrounds, written in Portugal and Spain during the Early Modern period and covering writers from different social background (cf. Marquilhas 2012; <http://ps.clul.ul.pt>).

Currently, one can also notice a trend (back) towards multi-genre corpora, combining handwritten ego-documents with printed data such as newspapers. Elspaß & Niehaus (2014: 51-52), for instance, suggest a corpus design for German which considers emigrant letters as historical data *from below* on the one hand, and regional newspapers as historical data *from above* on the other. Furthermore, an intriguing corpus project on nineteenth-century Icelandic is currently being compiled at the University of Iceland, comprising private letters, school assignments and newspapers (cf. <http://www.arnastofnun.is/page/LCLV19>; cf. also van der Feest Viðarsson 2017).

Specifically for the research presented in this dissertation, a new multi-genre corpus of late eighteenth- and early nineteenth-century Dutch was compiled. The *Going Dutch Corpus* contains (handwritten) ego-documents, viz. private letters, diaries and travelogues, as well as (printed) newspapers. Two diachronic cross-sections, viz. 1770–1790 and 1820–1840, allow a systematic analysis of the effectiveness of the official language policy introduced in the Northern Netherlands in 1804/1805. The texts in the corpus cover seven regions, both from the centre and the periphery, and in the case of ego-documents, represent both male and female writers. The design and compilation of the *Going Dutch Corpus* will be presented at length in Chapter 4.

5 Prescriptivism, language planning and policy

Among the variety of issues that have been addressed in historical sociolinguistics are the interrelated topics of linguistic prescriptivism, language planning and policy, all of which are "examples of endeavors to influence the way that languages are used within a speech community" (Auer et al. 2015: 8). Often with a strong focus on prescriptivism, a great deal of attention in historical-(socio)linguistic research has been paid to the metalinguistic comments made by grammarians and norm authorities (e.g. Tieken-Boon van Ostade 2008; cf. also Auer & Gonzáles-Díaz 2005: 31). More recently, however, there has been growing interest in the interplay of prescriptivism and normativity on the one hand, and actual language practices on the other (e.g. Rutten et al. 2014a; Anderwald 2016; cf. also Ziegler 2007).

While the study of the relationship between language norms and language usage is not new as such, the advances in (corpus) linguistics in recent decades, notably the field of historical sociolinguistics, have made it possible to reassess the interplay of norms and usage through the systematic study of large corpora. The edited volume by Rutten et al. (2014a), in fact, discusses language norms and language use in the histories of Dutch, English, French and German from the seventeenth to the nineteenth centuries, in most cases on the basis of corpus studies. What has widely been neglected, though, is the potential influence of language planning and language policy in historical contexts. Up to today, it is very much unclear whether and to what extent these external factors have affected linguistic patterns of variation and change.

Language planning is commonly divided into three major parts: status planning, corpus planning, and acquisition planning[16]. Wright (2012) embeds these three types of language planning in the context of nation-building and nationalism. According to her definition, status planning is "the process whereby state elites identify a language variety as the national language to be used in all the formal functions of state business" (Wright 2012: 65). The conscious activity of corpus planning, which includes, for instance, the codification of linguistic norms in orthographies and grammars, typically evolves from the "growing desire to achieve and maintain linguistic cohesion". Therefore, it also plays an essential role in the process of (linguistic) unification within the nation (Wright 2012: 68). Finally, acquisition planning aims at the promotion of the national language through the educational system. Wright (2012: 71) points out that school was "the institution where the ideology of one people, one territory and one language could be translated into reality", leading to a generation of children who become literate in the national standard language.

Terminologically differentiated from language planning, language policy, according to Ager (2001: 5-6), is "official planning, carried out by those in political authority, and has clear similarities with any other form of public policy". In that sense, language policy "represents the exercise of political power, and like any other policy, may be successful or not in achieving its aims" (ibid.). In the Dutch case, as examined in this dissertation, various language planning activities led to a concrete national language policy in the early nineteenth century on behalf of the Batavian government (cf. Chapter 2). The policy measures comprised the publication of an officialised orthography (Siegenbeek 1804) and grammar (Weiland 1805) for the

[16] Ager (2001: 6) provides broad definitions of the three types: "Status planning modifies the status, and hence the prestige, of languages or language varieties within society, often by modifying the way the language codes groups or individuals use are perceived ('deliberate efforts to influence the allocation of functions among a community's languages'). Corpus planning in the traditional sense is what communities do to the forms of the language ('graphization, standardisation, modernization and renovation'), but is sometimes also subdivided into codification of the existing language together with its elaboration and modernization by adding new terms or styles and controlling neologisms. Acquisition planning affects the 'acquisition, reacquisition or maintenance of first, second or foreign languages'" (cf. also Cooper 1989).

educational and administrative domains. Investigating the effectiveness of this *schrijftaalregeling* 'written language regulation', I therefore use the term language policy (rather than planning) in order to refer to the top-down endeavour of the government in the early 1800s to exert influence on actual language usage.

The 'success' of concrete language planning and/or policy measures, but also of prescriptivism in a more general sense, has hardly been examined in historical-(socio)linguistic research. Empirical studies based on large-scale corpora are needed for a systematic assessment of how planning and policy measures affect usage patterns. There have been only a few efforts to test their effectiveness quantitatively. The study by Auer & González-Díaz (2005), for example, aims to reassess the effects of eighteenth-century prescriptivism on actual language use in English by closely comparing two corpora: a so-called precept corpus (i.e. a collection of metalinguistic comments) and a usage corpus based on data from the *Helsinki Corpus of English Texts* and *ARCHER* (cf. Section 4). This method of assessing the influence of prescriptive grammars on actual usage had been introduced in studies on German standardisation (e.g. Takada 1998; Langer 2001).

A more recent empirical study investigating the success of prescriptivism is presented in Anderwald (2016). Focusing on verbs and verb categories in nineteenth-century English, Anderwald's study correlates the linguistic norms laid down in British and American English grammars with a corpus-based analysis of actual language change. For the French case, the article by Poplack et al. (2015) contributes to the discussion on prescriptivism and usage by studying the evolution of morphosyntactic features (such as the future temporal reference, cf. also Poplack & Dion 2009) in the grammatical tradition and a corpus of usage data.

Addressing the possible influence of prescriptive norms on language usage in the history of Dutch, Nobels & Rutten's (2014) paper on seventeenth-century Dutch studies two features, viz. negation and the genitive case. For the eighteenth century, Simons & Rutten (2014) examine the representation of final *n* and, again, the genitive case. Both studies, which are based on data from the *Letters as Loot* corpus (cf. Section 4), show that there is only limited evidence that language users adhered to the prescriptive norms codified in normative publications. The seventeenth-century data, for instance, show that bipartite negation, against its stigmatisation in metalinguistic discourse, remained a common and fairly frequent variant across large parts of the population. In contrast, the eighteenth-century case study on the deletion of final *n* reveal a clear influence of norms (in favour of <en>) on usage. Nevertheless, it should be emphasised that both periods under investigation in Nobels & Rutten (2014) and Simons & Rutten (2014), i.e. the late seventeenth and late eighteenth centuries, respectively, were before the introduction of the first Dutch language policy in the early 1800s – and thus before the standardisation of Dutch became a national concern and was regulated on behalf of the government. In fact, hardly anything is known about the influence of these official language regulations and their possible normative influence on actual usage patterns.

The aim of my dissertation is to fill this research gap by investigating the effectiveness of the concrete language policy measures in the Northern

Netherlands. As a top-down policy, the introduction of the Dutch *schrijftaalregeling* and its implementation into the national school system have to be regarded as a decisive intervention in the sociolinguistic situation of the early nineteenth century (cf. Chapter 1). The extent to which these changes affected language practices will be examined based on a newly compiled multi-genre corpus of late eighteenth- and early nineteenth century Dutch (cf. Chapter 4 for a detailed outline of the *Going Dutch Corpus*). More generally, my dissertation also aims to contribute to the broader historical-sociolinguistic discussion on the impact of language planning and policy in the past (and how to measure it quantitatively). Not only in the history of Dutch, but in most European language histories, this has been an astonishingly understudied topic of research.