Letters as Loot
A sociolinguistic approach to seventeenth- and eighteenth-century Dutch

Gijsbert Rutten
and Marijke J. van der Wal

John Benjamins Publishing Company
Letters as Loot
Over the last three decades, historical sociolinguistics has developed into a mature and challenging field of study that focuses on language users and language use in the past. The social motivation of linguistic variation and change continues at the forefront of the historical sociolinguistic enquiry, but current research does not stop there. It extends from social and regional variation in language use to its various communicative contexts, registers and genres, and includes issues in language attitudes, policies and ideologies. One of the main stimuli for the field comes from new digitized resources and large text corpora, which enable the study of a much wider social coverage than before. Historical sociolinguists use variationist and dialectological research tools and techniques, perform pragmatic and social network analyses, and adopt innovative approaches from other disciplines. The series publishes monographs and thematic volumes, in English, on different languages and topics that contribute to our understanding of the relations between the individual, language and society in the past.

For an overview of all books published in this series, please see

www.benjamins.com/catalog/ahs

Editors

Marijke J. van der Wal
Leiden University

Terttu Nevalainen
University of Helsinki

Editorial Board

Wendy Ayres-Bennett
University of Cambridge

Martin Durrell
University of Manchester

Agnieszka Kielkiewicz-Janowiak
Adam Mickiewicz University, Poznań

William A. Kretzschmar Jr.
University of Georgia, Athens GA

Mieko Ogura
Tsurumi University, Yokohama

Suzanne Romaine
University of Oxford

Daniel Schreier
University of Zurich

Merja Stenroos
University of Stavanger

Sali A. Tagliamonte
University of Toronto

Ingrid Tieken-Boon van Ostade
Leiden University

Donald N. Tuten
Emory University, Atlanta GA

Wim Vandenbussche
Vrije Universiteit Brussel

Anna Verschik
Tallinn University

Volume 2

Letters as Loot.
A sociolinguistic approach to seventeenth- and eighteenth-century Dutch
by Gijsbert Rutten and Marijke J. van der Wal
Letters as Loot

A sociolinguistic approach to seventeenth- and eighteenth-century Dutch

Gijsbert Rutten
Marijke J. van der Wal
Leiden University
Table of contents

Preface and acknowledgements xi

CHAPTER 1
Letters as Loot: A historical-sociolinguistic challenge 1
1. The Letters as Loot source 1
2. Letters as Loot and language history ‘from below’ 2
3. Selection of periods and phenomena 5
4. Speech and writing 7
5. The variables 9
6. The autograph status of letters, corpora and presentation of results 13
7. Outlook 17

CHAPTER 2
Sounds and signs: From local to supralocal usage 19
1. Introduction 19
2. Standard languages, language standards and the degree of orality 19
3. The degree of orality in the history of Dutch 21
4. Case studies Zeeland 23
   4.1 The subcorpora used 23
   4.2 H-dropping in the seventeenth century: First case study 24
   4.3 H-dropping in the seventeenth century: Second case study 26
      4.3.1 Deletion of h 26
      4.3.2 Prothesis of h 30
      4.3.3 Substitution of ⟨h⟩ for ⟨a⟩ 31
   4.4 H-dropping in the eighteenth century, and diachronically 32
   4.5 Long e’s in the seventeenth century 34
   4.6 Long e’s in the eighteenth century, and diachronically 41
   4.7 Conclusions 44
5. Case studies North Holland and Amsterdam 46
   5.1 The subcorpora used 47
   5.2 Germanic sk 49
      5.2.1 Seventeenth-century North Holland 50
      5.2.2 Seventeenth-century Amsterdam 51
      5.2.3 Eighteenth-century North Holland 52
      5.2.4 Eighteenth-century Amsterdam 53
5.3 Germanic ft 54
  5.3.1 Seventeenth-century North Holland 55
  5.3.2 Seventeenth-century Amsterdam 56
  5.3.3 Eighteenth-century North Holland 57
  5.3.4 Eighteenth-century Amsterdam 58
5.4 A-like vowels in the seventeenth century 58
  5.4.1 The representation of a-like vowels 58
  5.4.2 The writing systems used 61
5.5 Prefix ge- in the seventeenth century 65
5.6 Long e’s in the seventeenth and eighteenth centuries 67
5.7 Conclusions 72
6. Conclusions 73

CHAPTER 3
Epistolary formulae: Functions and text composition 75
1. Introduction 75
2. A sample letter 77
3. The subcorpora 81
4. Functions of epistolary formulae 81
   Text-constitutive formulae 82
   Intersubjective formulae 83
   Christian-ritual formulae 83
5. Text composition 85
6. Text-constitutive formulae 86
   6.1 Text-type formulae 86
      6.1.1 Address 87
         Prepositional phrases having the meaning “to” 87
         The addressee’s name and social role 88
         The addressee’s location 90
         The means by which the letter is sent 92
         Additional information 93
      6.1.2 Date 93
         Praise to God 94
         The writer’s location 94
         The date 95
         Additional information 97
      6.1.3 Opening 97
         Addressing the recipient 98
         Greeting the recipient 100
         Referring to (earlier) communication 101
      6.1.4 Closing 105
   6.2 Text-structural formulae 108
      6.2.1 Initiating discourse 108
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.2.2</td>
<td>Continuing discourse</td>
<td>109</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Ending discourse</td>
<td>111</td>
</tr>
<tr>
<td>7.</td>
<td>Intersubjective formulae</td>
<td>114</td>
</tr>
<tr>
<td>7.1</td>
<td>Health formulae</td>
<td>114</td>
</tr>
<tr>
<td>7.1.1</td>
<td>Health statements</td>
<td>114</td>
</tr>
<tr>
<td>7.1.2</td>
<td>Health wishes</td>
<td>117</td>
</tr>
<tr>
<td>7.1.3</td>
<td>Subordinate health formulae</td>
<td>118</td>
</tr>
<tr>
<td>7.1.4</td>
<td>Health formulae in the eighteenth century</td>
<td>121</td>
</tr>
<tr>
<td>7.2</td>
<td>Greeting formulae</td>
<td>121</td>
</tr>
<tr>
<td>7.3</td>
<td>Contact formulae</td>
<td>123</td>
</tr>
<tr>
<td>8.</td>
<td>Christian-ritual formulae</td>
<td>125</td>
</tr>
<tr>
<td>9.</td>
<td>Conclusions</td>
<td>127</td>
</tr>
</tbody>
</table>

CHAPTER 4

Variation and change in formulaic language | 129
1. Introduction | 129
2. Formulaic language and writing experience | 130
   2.1 Previous research | 130
   2.2 Processing effort and writing experience | 131
   2.3 Writing experience and social variables | 133
   2.4 Summing up – hypotheses | 134
3. Case studies | 135
   3.1 The subcorpora | 136
   3.2 Greeting | 137
   3.3 Communicating health | 141
   3.4 Communicating sorrow | 146
   3.5 Communicating God’s omniscience | 150
   3.6 Initiating discourse | 153
   3.7 Continuing discourse | 158
   3.8 Conclusions | 163
4. Variation and change as functions of writing experience | 164
   4.1 The social literacy boundary | 164
   4.2 Gender differences | 167
   4.3 Incremental and levelling writing experience | 169
   4.4 Writing experience vs. social identity and style | 170
5. Conclusions and further discussion | 171

CHAPTER 5

Detailing the writing process: Formulaic language, social and professional writers, and the influence of letter-writing manuals | 173
1. Introduction | 173
2. The non-autograph subcorpus | 174
3. Comparing autographs and non-autographs 174
   3.1 Greeting 174
   3.2 Communicating health 175
   3.3 Communicating sorrow 176
   3.4 Communicating God’s omniscience 177
   3.5 Initiating discourse 177
   3.6 Continuing discourse 178
   3.7 Conclusion 179
4. Writing in the name of: Social or professional writers 180
5. Formulaic language, group practices and social routine 185
6. Formulae in letter-writing manuals 187
   6.1 Elite manuals 188
   6.2 School books 190
   6.3 Jacobi and his successors 195
7. Conclusions 202

CHAPTER 6
Forms of address 203
1. Introduction 203
2. Forms of address in the history of Dutch 204
3. Politeness as a variable 207
   3.1 Private vs business 208
   3.2 Within the core family 209
   3.3 The j-forms 213
      3.3.1 The seventeenth century 213
      3.3.2 The eighteenth century 218
   3.4 An increase in negative politeness? 222
   3.5 Politic behaviour and social stratification 223
4. Sociolinguistic variation and change 225
   4.1 The overall diachronic picture 226
   4.2 Epistolary forms of address 227
      4.2.1 The seventeenth century 227
      4.2.2 The eighteenth century 231
      4.2.3 Social effects on the change from ul to ue 235
      4.2.4 The contexts in which epistolary forms appear 236
   4.3 The pronominal forms of address 239
5. Politeness – again 243
6. Conclusions 245

CHAPTER 7
Clause chaining between spoken and written language 247
1. Introduction 247
2. The subcorpora 250
3. Clause linking in the letter corpus 251
   3.1 No bleached connectives and no punctuation 252
   3.2 Bleached connectives, but no punctuation 254
   3.3 Bleached connectives as well as punctuation 255
   3.4 Punctuation, but no bleached connectives 256
4. Grammaticalisation of clause linkages and discourse markers 257
5. The rise of punctuation 266
6. Variation and change in clause chaining 269
   6.1 Discourse units 271
   6.2 Sociolinguistic patterns 280
7. The interplay of punctuation and clause chaining 284
8. Conclusions 286

CHAPTER 8
Variation and change in the relative clause 289
1. Introduction 289
2. Relativisation in the history of Dutch 289
3. Syntacticisation of the relative clause – hypotheses 293
4. Variation and change in the relative clause – case studies 296
   4.1 Extracting ¿s and ¿s – diachronic results 296
   4.2 Social and gender variation 297
   4.3 A change from above 302
5. Relativisation and incremental writing experience 302
6. Relativisation and formulaic language 304
   6.1 Epistolary formulae 305
   6.2 Constructional diffusion 313
7. Relativisation and clause integration 313
   7.1 Integration of relative clauses 314
   7.2 Integration and syntacticisation 318
8. Conclusions 320

CHAPTER 9
Apocope of final schwa 323
1. Introduction 323
2. The subcorpora 325
3. A regionally diffused phonological change 326
   3.1 The broad picture by time and region 326
   3.2 Right context 331
   3.3 Left context 335
4. North Holland, including Amsterdam 338
5. South Holland 341
6. Zeeland 344
Letters as Loot

7. Interpreting social and regional variation 347
8. Epistolary conventions 350
   8.1 Formulaic language 350
   8.2 Zero subjects 355
9. Conclusions 360

CHAPTER 10
Clausal and local negation 363
1. Introduction 363
2. The subcorpora and the overall results 364
3. Regional patterns 365
4. Constructional patterns 368
5. Lexical-semantic patterns 373
   5.1 The negators niet and geen 373
   5.2 Type of verb 374
6. Phonological patterns 378
7. Complexity effects 380
   7.1 Proximity 381
   7.2 Verbal cluster complexity 383
8. Social patterns 385
9. Social patterns in Amsterdam and Zeeland 388
10. Conclusions 391

CHAPTER 11
Harvesting: Reflection and evaluation 393
1. Harvesting: Introduction 393
2. Mapping variation and change 393
3. Explanations: Community practices and politeness theory 396
4. The writing process and different degrees of writing experience:
   Connecting all the data 398
   4.1 Acquiring writing skills 398
   4.2 Phenomena and the explanation of writing experience 401
5. The complex nature of change 404
6. The fruits of confiscated letters 405
7. Conclusions 408

References 409
Index 425
Preface and acknowledgements

The present volume is the fruit of the Letters as Loot. Towards a non-standard view on the history of Dutch research programme (2008–2013), funded by the Netherlands Organisation for Scientific Research (NWO) and conducted at the Leiden University Centre for Linguistics. In this programme, initiated and directed by Marijke van der Wal, Gijsbert Rutten participated as a postdoctoral researcher. We both would like to acknowledge our gratitude to a number of persons and institutions which were most helpful and supportive in various stages of our research project.

Our research would not have been possible without the facilities provided by the National Archives (Kew, UK) where the impressive source of confiscated Dutch documents is currently kept. It was always a great pleasure for our research team to work in the Map and Large Document Reading Room, discovering new and surprising finds in the High Court of Admiralty’s (HCA) archive. Thanks are also due for giving us the opportunity to make digital photos during our frequent visits. With a few exceptions, all photos of the original documents represented in the present volume were taken – courtesy of the National Archives – by members of our Letters as Loot research team. Three photos (Figures 2, 3 and 4 in Chapter 5) were provided by the Metamorfoze project, a national enterprise aimed at preserving Dutch paper heritage.

The way for our archival research was paved by historians Els van Eyck van Heslinga and Perry Moree, both at the time affiliated to the National Library of the Netherlands (The Hague), who over the years continued to stress the importance of the ‘sailing letters’, and by Roelof van Gelder (Amsterdam) who made the indispensable inventory of more than one thousand HCA boxes. Erik van der Doe (National Archive, The Hague) and Dirk Tang (National Library of the Netherlands) were helpful in various respects related to the Metamorfoze project.

The Leids Universiteitsfonds (LUF) supported a pilot project carried out by Marijke van der Wal in order to establish whether the private letters among the HCA documents were indeed such a treasure for historical linguists as she suspected them to be. Once it became clear that this was indeed the case, the contributions by participants of the Wikiscripta Neerlandica project, which ran successfully from 2007 till the end of 2011, were invaluable for our research. The volunteers of this Leiden-based project, initiated by Marijke van der Wal, carried out most of the time-consuming transcription process and at many occasions showed their enthusiasm for the extraordinary letters. We are greatly indebted to them all. Thanks are also due to Coen Zimmerman (Leiden) who constructed a tailor-made database to store and retrieve all the information on the letters in our corpus.
Special and heartfelt thanks go to the members of our Letters as Loot team who shared our fascination for the rediscovered letters: the PhD-researchers Judith Nobels and Tanja Simons, who both successfully defended their theses *(Extra)Ordinary letters* and *Ongekend 18e-eeuws Nederlands* in 2013, and the research assistants Reinald Molenaar and Juliette Sandberg, who were such a great help in administrative, practical and transcription matters. Students, both at BA and MA level, were involved in our project, of whom we would like to mention Chrisje Meima in particular. From the very start Elisabeth Koning (Oegstgeest) has been a most inspiring companion who took thousands of photos during our visits to the National Archives and who also gave her invaluable advice on many issues over the years.

We gratefully acknowledge the contributions of various colleagues of the Institute for Dutch Lexicology (INL, Leiden): Marijke Mooijaart and her lexicological assistants who carefully tagged and lemmatised our corpus, and Katrien Depuydt and Jesse de Does who took care of the advanced technology needed for a user-friendly application. This excellent collaboration resulted in making available our whole corpus, including photos, transcriptions and metadata, through an attractive website provided with elaborate search tools (http://brievenalsbuit.inl.nl).

During our long-term project many colleagues and others took a great interest in our letter treasure: colleagues in the Department of Dutch Language and Culture, in the Leiden University Centre for Linguistics, in other parts of the Leiden Faculty of Humanities and also linguists abroad. We cannot list them all individually, but we would like to mention Wim van den Doel, the Dean of the Faculty of Humanities, Ton van Haaften, the director of LUCL, and Ingrid Tieken-Boon van Ostade, who has a chair in English Socio-historical Linguistics: they all showed their continued interest in our research, stressing the impact of this important project. And last but not least we would like to mention with the utmost gratitude the continuous support and encouragement that we received from the Merweborg foundation, in particular from the members of its board of curators Menno Witteveen (The Hague), Judith Pollmann (Leiden) and Jos Schaeken (Leiden/The Hague).

Over the years we were both heavily involved in the Letters as Loot research, which led to a range of publications and to draft chapters for the present volume. During the last stage of the project Marijke van der Wal transformed the draft texts, mainly based on research by Gijsbert Rutten, into the present book, which was subsequently accepted for the AHS series. We thank Marilyn Hedges (Leiden/The Hague) for her invaluable linguistic advice and for carefully correcting our final English texts. During the peer review process that followed we benefited greatly from the constructive comments made by the anonymous referees for John Benjamins Publishing Company. We also highly appreciate the editorial advice given by Terttu Nevalainen, one of the editors of the AHS series, and Anke de Looper (John Benjamins Publishing Company) during the various stages of the publishing process.
Finally, for us both *Letters as Loot* marks the end of a period of inspiring collaboration concentrated on a unique source of letters from the past. We hope that the readers of our book will discover the immense value of this intriguing source which we expect will continue to inspire historical linguists and scholars from various other disciplines.

Gijsbert Rutten and Marijke van der Wal
Leiden, April 2014
1. The Letters as Loot source

It is something of a once-in-a-lifetime experience for a historical linguist to immediately understand the far-reaching perspectives opened up by a newly rediscovered text source. This is what the second author of the present book experienced when she was presented with the still rough inventory made of the so-called ‘sailing letters’, an impressive source of documents, currently kept in the British National Archives (Kew, UK; van Gelder 2006). These documents are remnants of times of warfare and privateering, the longstanding legitimate activity of conquering enemy ships, which in the past was performed by all seafaring European countries and regulated by strict rules.

The ‘sailing letters’ comprise approximately 40,000 Dutch letters that, with all other papers, both commercial and private, were confiscated aboard ships taken by the English fleet and by private ships (privateers) during the frequent warfare between England and the Netherlands, from the second half of the seventeenth century to the early nineteenth century (van Gelder 2006: 31). Among these are some 15,000 private letters. To fully appreciate the huge number of letters it is important to note that in very many cases the ships’ cargo contained far more mail than the crew’s own correspondence and the letters that the crew had in their possession to deliver for relatives, friends and acquaintances.1 Ships often took mailbags on board, thus functioning as mail carriers between the Netherlands and regions such as the Caribbean, western Africa and Asia.

In the legal procedure that followed the taking of ships in England, the confiscated papers were considered evidence in cases where the High Court of Admiralty (HCA) had to decide whether the conquered ships were taken lawfully. If the rules had been followed scrupulously, the ship and all its cargo were considered as loot for the privateer (van Gelder 2006: 10–15; van Vliet 2007: 53–55). Once the final decision had been reached, the papers, which, apart from the letters, comprise a wide range of other material including treatises on seamanship, plantation accounts, textile samples, ships’ journals, poems and lists of slaves, remained in the HCA’s Archives. Miraculously, they have survived to the present day, and are currently stored in over a thousand boxes in the British National Archives.

1. Many letters were obviously sent through friends according to the phrase met vryent dye god gheleyde “with a friend whom God may lead” and its variants on the outside of the letters. This method of delivering mail was common practice in the seventeenth and eighteenth centuries.
A brief pilot study by Marijke van der Wal showed the uniqueness of the source material by confirming the assumption that the private letters stem from a rich variety of senders, both men and women (and even children) of all social ranks. The material proved to include letters from the lower classes which have otherwise left very few written traces (van der Wal 2006: 9–11; 2008: 95–99). This text source, therefore, called for a historical-sociolinguistic approach, more in particular what has become known as a ‘language history from below’ approach (Elspaß e.a. 2007; Elspaß 2012a: 160–162). This conclusion led to the start of the Letters as Loot research programme, funded by the Netherlands Organisation for Scientific Research (NWO) and conducted at the Leiden University Centre for Linguistics (LUCL).2 The present monograph is the fruit of joint research within this Letters as Loot programme.

2. **Letters as Loot and language history ‘from below’**

The history of Dutch is very similar to that of other western European languages in that it has been dominated by the standardisation perspective for a considerable time.

---

The Dutch standardisation process included the familiar development from mainly
dialect diversity in the Middle Ages towards a supraregional standard language in sub-
sequent centuries. Codification started in the second half of the sixteenth century and
has continued with the publication of grammars, dictionaries and other normative
texts ever since. In the standardisation process, the region of Holland took the lead as
the political, economic and cultural superpower. The traditional view of the stand-
ardisation of Dutch is largely based on the language of printed texts that were mainly
written by well-educated upper-class men. Over the centuries, the written language of
this small upper layer of society became increasingly uniform, which has given the
impression of a standard language gaining more and more ground. Whereas the sev-
enteenth-century printed language still allowed a fair degree of variation, the eight-
teenth-century printed language shows considerable uniformity. Linguistic uniformity
was therefore assumed to have been consolidated in the eighteenth century. This type
of traditional language history, which has been characterised as a form of tunnel vision
(Watts & Trudgill 2002), relied on the language use of the highest ranks in society,
which, according to Elspaß (2007: 4–5), consisted of no more than five per cent of the
contemporary population.

An approach such as that from the perspective of ‘below’ has a much broader scope,
and explicitly focuses on the language of the middle and lower classes and their hand-
written texts (Elspaß 2012a: 160–162). However, despite the increasing awareness that
the standardisation process and the ‘remaining’ linguistic diversity are two sides of the same coin that should be addressed equally and should complement each other, practical problems such as finding appropriate data may hinder such a broader approach. Until recently, apart from studying printed texts, linguists had to rely mainly on private documents written by men from the higher ranks of society for the history of the Dutch language. Private documents from women in general and from both men and women of the lower and middle classes were available only in very small numbers, scattered over various archives in the Netherlands. Against the background of these scarce data, the importance of the ‘sailing letters’ source stands out. Since the confiscated private letters were shown to compensate for this lack of linguistic data, we were able to explore, in the Letters as Loot research programme, what is largely a terra incognita of Dutch language history from below. In other words, the letters offered an excellent opportunity for mapping linguistic diversity and filling in the gaps left by traditional historical linguistics that in many cases had a teleological perspective on language history, as they focused mainly on literary texts and formal texts from higher registers (van der Wal 2006).

Our large-scale exploration of private letters took place within a historical sociolinguistic framework that has broadened its scope over the past three decades. Multifaceted historical-sociolinguistic research in different language areas and for different time periods (cf. Hernández-Campoy & Conde-Silvestre 2012) comprises different trends of research. The research ranges from studies of genre variation by multi-genre corpora such as the Helsinki Corpus of English Texts, regional variation, for instance in Elspaß’s corpus of nineteenth-century German emigrant letters (Elspaß 2005, 2012b: 47–48), and social, gender and age variation in corpora provided with metadata on the writers of the corpus texts, such as the Corpus of Early English Correspondence (cf. Nevalainen & Raumolin-Brunberg 1996, 2003) to social network studies at the core of which are individuals and their relationships (e.g. Tieken-Boon van Ostade 2005). It is no coincidence that many of the corpora explored for this research consist of private letters (cf. Elspaß 2012a: 162–163 for an overview of such corpora). Of all written text types, private letters are considered to be as close to speech as non-fictional historical texts can possibly be, which means that such letters allow us to trace the informal language and linguistic variation of the past (van der Wal 2006; Elspaß 2012a: 156).

More in particular, our research tied in with an approach from the perspective of language history from below, which focuses on non-literary everyday language, presumably found in ego-documents such as letters and diaries from lower- and

---


4. For Dutch, Vandenbussche and Willemyns and their research group at the Vrije Universiteit Brussel (VUB) have concentrated on historical-sociolinguistic research of mainly administrative text types of the nineteenth-century Southern Netherlands (cf. Willemyns & Vandenbussche 2000; Vandenbussche 2002; Vosters & Vandenbussche 2012).
middle-class writers (Elspaß 2012a: 160–162; van der Wal & Rutten 2013c). To avoid any misunderstanding, we stress that the term ‘from below’ primarily refers 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 (Elspaß 2007: 5). The ambiguous and sometimes confusing term ‘from below’ will not be used to indicate only linguistic phenomena from below the level of awareness (Labov 1994: 78) or linguistic changes that start in the lower social classes. In the present book, we take into account various kinds of linguistic phenomena, some below and others above the level of awareness, and various kinds of changes that appear to have spread either from below or from above socially. Generally speaking, we consider making available insights into the as yet mainly unknown written (and spoken) language of major parts of the seventeenth- and eighteenth-century Dutch population, men and women alike, to be one of the challenges of our research. To further detail our goals, we aim to map variation and change (1), present explanations for the variation and change found (2), fill gaps in the history of Dutch (3) and contribute to the international historical-sociolinguistic debate (4). In the final chapter of this volume, evaluating both the epistolary material and the historical-sociolinguistic approach applied, we will assess to what extent we succeeded in achieving these goals. In the present chapter, we discuss the main issues of our research and the choices that had to be made prior to conducting the research.

3. Selection of periods and phenomena

Before starting to analyse the letters from a historical-sociolinguistic perspective, a selection had to be made from the various periods when England and the Netherlands were at war with one another: the First Anglo-Dutch War (1652–1654), Second Anglo-Dutch War (1665–1667), Third Anglo-Dutch War (1672–1674), Fourth Anglo-Dutch War and the American War of Independence (1776–1784), and the Napoleonic period (1793–1813).5

Two periods of about ten years each were chosen for making cross-sections in the approximately 15,000 private letters: the seventeenth-century period of 1664–1674, from the prelude to the Second to the end of the Third Anglo-Dutch War, and the eighteenth-century period of 1776–1784, from the start of the Fourth Anglo-Dutch War to the end of the American War of Independence.6 Letters from these different

5. Apart from these periods, documents survived from ships taken during the War of the Austrian Succession (1739–1748) and the Seven Years’ War (1756–1763), periods of partial neutrality and changing alliances (van Gelder 2006: 15, 30).

periods, with an intended interval of approximately one hundred years, were selected and digitally photographed during a number of visits to the National Archives. From the digital photographs taken, transcriptions of the letters were made following a diplomatic transcription method without any normalisation of word boundaries or of i/j and u/v variation. Thus we transcribed *ghe coft*, *iaer* and *bouen* which we did not change into *ghecoft* “bought”, *jaer* “year” and *boven* “above” as will become apparent in the examples and quotations in the following chapters. We opted for this transcription method in order to enable various kinds of linguistic analyses, including phonological and graphematic analyses.

The letter corpora we compiled, the detailed structure of which we will discuss below, allow for studies at all linguistic levels, from phonology and orthography to morphosyntax and text composition. For the present monograph, we concentrated on

---

7. The time-consuming transcription process was carried out mainly by volunteers of the *Wikiscripta Neerlandica* project, which was initiated by Marijke van der Wal and ran successfully from 2007 until the end of 2011. The participants provided transcriptions that were carefully checked by different members of the *Letters as Loot* research team before they were accepted as final transcriptions for the electronic corpus.

8. Both the transcriptions and the corresponding digital photographs were included in a digital archive. As a working tool, a database, specifically developed for the *Letters as Loot* research programme, was constructed to store and retrieve all the information on the letters. The transcriptions, digital photographs and metadata of our final research corpus (cf. Table 2) are available at http://brievenalsbuit.inl.nl.
two types of phenomena at different linguistic levels: phenomena characteristic of the
text type of letters and frequent phenomena that have been the topic of debate in Dutch
historical linguistics, which two categories may obviously coincide. The often assumed
relative proximity of private letters to spoken language (cf. Section 4 below), should not
obscure the value of written language in its own right, a value that has been stressed by
such authors as Jucker & Taavitsainen (2003: 7) and Martineau (2013). The latter points
out the hybridity of letters and other ego-documents, in that they contain both typi-
cally oral elements and linguistic features characteristic of the written code. As both
oral and written characteristics warrant the attention of historical linguists, the most
striking epistolary characteristic of our letter corpus, the frequent occurrence of for-
mulaic language, requires a detailed study, which will be presented in Chapters 3, 4 and
5. These chapters focus on the different functions of epistolary formulae and their role
at the level of text composition (Chapter 3) and examine variation and change in the
use of formulaic language (Chapter 4). The actual practice of different letter writers,
both in self-written and non-self-written letters, and the much-questioned role of let-
ter-writing manuals are the topic of Chapter 5. Examining formulaic language, we also
address one of the main questions of the present monograph, viz. what happened when
people from various backgrounds put pen to paper and started to write. Applied to this
specific case, did they rely on formulaic language and, if they did, to what extent? Other
major questions that will be dealt with in this context include whether the variation in
the use of formulaic language is sociolinguistically patterned, and whether the use of
formulae changed from the seventeenth to the eighteenth century. These questions ob-
viously relate to our main goals of mapping and explaining variation and change.

Forms of address, which comprise both variants characteristic of the epistolary
context and general variants, are among the other morphosyntactic phenomena se-
lected that occur both in speech and in writing. The chosen features of relative clauses,
apocope of final -e and negation are relatively highly frequent phenomena in our letter
corpora, and were all undergoing change in seventeenth- and eighteenth-century
Dutch. As our sociolinguistically stratified letter corpora offer new ways of analysing
these changes, our main aim is to establish possible sociolinguistic patterns in the
spread of these changing phenomena. Other phenomena selected for the present
monograph can be related to one of the main historical-sociolinguistic issues of de-
bate, that is the relationship between speech and writing in documents of the past.

4. Speech and writing

Relying on earlier studies in historical sociolinguistics, in which informal private let-
ters were found to be close to the oral modes of expression and therefore were as-
sumed to give access to the everyday spoken language of their writers, we may expect
our confiscated letters to reveal speech features at different linguistic levels. On closer
examination, the letters do contain traces of the contemporary spoken language,
phonological, morphological as well as syntactic. We find, for instance, orthographies such as *maid* “maid” and *markte* “markets”, reflecting the alternative pronunciation of *[ei]* as *[ai]* and the deletion of final *n* in speech, the alternative diminutive suffix *-ie* in *huysie* “little house”, and the syntactic *noun + possessive pronoun + noun* pattern to express the possessive such as in *een brief van de captaijn zijn vrouw* “a letter from the captain his wife/ the captain’s wife” (van der Wal 2007: 88–91, 93–95). These traces, which do not occur in most of the printed contemporary works, provide a window on the spoken language used in everyday communication. While, in this respect, the letters are considered to be of inestimable value, they also raise the questions of how spoken and written language are intertwined and to what extent these traces of the spoken language occur. From the perspective of local or dialectal usage, the question could be raised of whether the language of the private letters can be characterised as primarily representing local, dialectal usage, or whether they also contain elements of possibly supralocal writing conventions. It is against the background of diverging opinions on the respective influence of local dialects and supralocal writing systems that we carried out the mainly phonological and orthographic case study “Sounds and signs” in Chapter 2.

We may, in other words, question the letters’ degree of orality (Rutten & van der Wal 2011). The orality of various text genres has been discussed widely, including by Koch & Oesterreicher (1985) who differentiate between the language *medium* (i.e. phonic or graphic) and spoken and written *conception*, and determine the position of a text on a conceptual continuum between the poles of communicative *immediacy* (orality, informality, lack of planning) and *distance* (literateness, formality, a higher degree of planning). Being intended for interaction, private letters clearly are close to the language of immediacy, even more so than diaries and travelogues (Nevalainen & Raumolin-Brunberg 2003: 29; Elspaß 2012a: 157–159). We should note that the orality related to the language of immediacy thus differs from the earlier mentioned concept of orality as reflecting dialectal usage. Orality in this broader, non-dialectal sense is the topic of a syntactic case study into the phenomenon of clause chaining in Chapter 7 that examines the way in which letter writers linked clauses in order to create larger pieces of discourse. Being aware of different strategies in use – some more oral, others prototypically written – in this chapter we will determine the extent to which clauses are chained by the relatively oral use of *ende* and *en* “and” in order to establish another degree of orality. The approach taken in Chapter 2 on the degree of orality in the field of orthography and phonology and that in Chapter 7 on syntax and text composition share the effort to locate the genre of seventeenth- and

---

9. Cf. Nevalainen & Raumolin-Brunberg (2012: 32) who argue that “in our experience of compiling and using the Corpus of Early English Correspondence (CEEC), personal correspondence provides the ‘next best thing’ to authentic spoken language and, even with its obvious limitations, makes it possible to extend the variationist paradigm into the more distant past.” Cf. Nevalainen & Raumolin-Brunberg (2012: 29) and the publications mentioned there, for further discussion of different categories of written language in relation to speech.
eighteenth-century private letters on an oral–literate continuum. The leading theme of ‘orality’ versus writing conventions will also be addressed when dealing with morphological phenomena, such as verbal inflection, in Chapter 9.

Before being able to determine the nature of the private letters on an oral–literate continuum and to map variation and change in our letter corpora, in Section 5 we discuss the variables used, and in Section 6 we resolve the problematic issues of social stratification and of the autograph status of texts.

5. The variables

In mapping variation and change, we focus primarily on social variation, taking into account the variables of rank, gender and age. Of these variables, gender is relatively unproblematic, since, in almost all cases, the gender of the sender can be deduced from his or her name in the letter. Determining a sender’s age and social background may, however, raise practical problems, as letters, in most cases, only provide us with the sender’s and addressee’s name and address. Additional information is needed and has to be found in various – mainly municipal – archives, which hold local registers of marriage, baptism or burial and registers of contemporary notaries. Distinguishing between the categories of younger than 30, 30 to 50 and older than 50, we inferred the relative age of senders from their stage in life as a young person, a parent or a grandparent in many cases where detailed information was missing. The variable of rank not only presented us with practical problems, this variable is in itself also more problematic, as we have to decide on both the social stratification of the past and the degree of social-stratificational detail needed for linguistic research (cf. Nevalainen & Raumolin-Brunberg 2003: 136–138). The former is a familiar issue for historians who study the seventeenth- and eighteenth-century Northern Netherlands, often referred to as the Republic of the Seven United Provinces. Despite all kinds of possible reservations, we deemed it feasible to rely on their social stratification for this time span, which consists of six social ranks, viz. the highest rank, three layers of bourgeoisie, and two lowest ranks as represented in Table 1 (cf. Frijhoff & Spies 1999: 188–190; Bruijn 2008: 16).

We should add that, although the stratification into six ranks is a present-day construct, it is not a completely modern ‘invention’, but is partly based on contemporary observations (cf. Frijhoff & Spies 1999: 189). In our Letters as Loot corpus, relying on the stratification by historians, we distinguish four ranks or classes,10 since the highest rank of nobility and non-noble ruling classes is scarcely represented in our material and neither is obviously the class of have-nots. Actually, the fifth and sixth classes of

10. It should be noted that we use the term class as a synonym for rank, not as the nineteenth-century notion of class.
Table 1. Social stratification of the seventeenth- and eighteenth-century Republic of the Seven United Provinces.

<table>
<thead>
<tr>
<th>Historians’ stratification</th>
<th>Letters as Loot corpus</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Nobility and the non-noble ruling classes</td>
<td>Upper Class/UC</td>
</tr>
<tr>
<td>2. Bourgeoisie, e.g. wealthy merchants, ship owners, academics, commissioned officers</td>
<td>Upper-Middle Class/UMC</td>
</tr>
<tr>
<td>3. Prosperous middle class, e.g. large storekeepers, non-commissioned officers, well-to-do farmers</td>
<td>Lower-Middle Class/LMC</td>
</tr>
<tr>
<td>4. Petty bourgeoisie, e.g. petty shopkeepers, small craftsmen, minor officials</td>
<td>Lower Class/ LC</td>
</tr>
<tr>
<td>5. Mass of wage workers, e.g. sailors, servants, soldiers</td>
<td></td>
</tr>
<tr>
<td>6. Have-nots, e.g. tramps, beggars, disabled</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 could be considered as one single class, as there was a frequent exchange of sailors, soldiers, etc. from the wage workers class to the class of have-nots and vice versa (Frijhoff & Spies 1999: 190). The missing highest rank does not present us with a serious problem, since data from this rank are well-known and have dominated traditional language history. This leaves us with a social stratification of four ranks or classes, viz. the upper class (UC), the upper-middle class (UMC), the lower-middle class (LMC) and the lower class (LC), a four-partite classification that we would like to stress is relative rather than absolute. To assign specific letter writers to social ranks we used a variety of criteria, the most important being the writer's profession or, in the case of women, the husband's profession. Where other information on status in society or level of education was available, this played an additional role. Despite its partly present-day construct and although, in some cases, the social class to which a specific letter writer was assigned could be a matter of some dispute, we argue that the ranks distinguished or the relative four-way division could be considered as a hypothesis that may or may not be confirmed by the results of our investigations.11

Apart from the three sociolinguistic variables, the regional origin of the writers was also taken into consideration. It should be noted that, historically, we need to distinguish between the Northern and the Southern Netherlands, roughly corresponding to the present-day Netherlands and Belgium, respectively. In the Northern Netherlands, the seventeenth century is usually considered the ‘Golden Age’, both economically and culturally. This prosperity involved a remarkably high degree of

---

11. Palander-Collin et al. (2009: 5) also argue that “status differences and societal contexts may be operationalised in various ways, and only research will tell which categories were linguistically relevant”. Bergs (2012: 87) stresses the interplay between historians and linguists: whereas linguists use the models provided by historians, historians may use the linguistic results for refining their models.
urbanisation in the western parts of the Northern Netherlands, especially when compared to other western European countries. The most important regions demographically were Holland and Zeeland, both on the western coast of the Northern Netherlands. Around the middle of the seventeenth century, about 400,000 (c. 20%) of the almost 2 million inhabitants of the Northern Netherlands lived in the ports of Holland and Zeeland, in cities such as Amsterdam, Rotterdam, Middelburg and Vlissingen (Frijhoff & Spies 1999: 154). The city of Amsterdam, of metropolitan size given its 175,000 inhabitants, occupied a special place in this highly urbanised environment. At the same time, the Republic as a whole boasted nineteen towns of more than 10,000 inhabitants as opposed to only eight in England, fourteen in the Southern Netherlands, forty-four in France and twenty-three in Germany. These towns with more than 10,000 inhabitants comprised 32% of the total population of the Northern Netherlands, whereas this proportion was 21% in the Southern Netherlands, 7% in France and 4.4% in Germany (Frijhoff & Spies 1999: 157–158). Over the next century and a half, the population remained fairly stable (c. 2 million), and the degree of urbanisation stayed remarkably high. By 1800, more than one-third of the population (37%) lived in towns of more than 2,500 inhabitants, as opposed to only 12% of the population in France, 17% in Germany and less than 30% in England (Kloek & Mijnhardt 2001: 38).

Almost all the confiscated letters were sent to and from the Northern Netherlands, more particularly to and from the main regions of Holland and Zeeland. When discussing regional variation in the present monograph, we distinguish the four major regions of Zeeland (with its main cities of Middelburg and Vlissingen), South Holland (with its main city of Rotterdam), North Holland – Amsterdam, and North Holland – Other. Letters were labelled with a regional code in accordance with the region to which we could relate the letter writer (for details see Nobels 2013: 28–34). For instance, as a rule both letters written by people in Zeeland and letters written and sent from abroad to relatives in Zeeland received the regional code ‘Zeeland’. Thus letter writers born and raised in the Netherlands, who had migrated, either temporarily or definitively abroad, were labelled with the code of their regional origin. However, letters from second generation migrants who were raised in Dutch speaking families abroad received the regional code ‘Other/abroad’. See Figures 4 and 5, which give a contemporary geographical map and an overview of the main dialect areas in the present-day Dutch language area, thereby indicating the areas of North Holland, South Holland and Zeeland.

It should be mentioned that the distinction between North Holland and South Holland does not represent a contemporary administrative division, since Holland

---

12. Our corpora used for the present monograph (see Table 2) also comprise a smaller number of letters from and to other regions such as Flanders (seventeenth century), Utrecht and Gelderland (eighteenth century).
was a single province at the time. Amsterdam is considered separately for geographical as well as demographic reasons. Geographically, the city of Amsterdam is located in the south of North Holland, separated from the northern parts of North Holland by water. Demographically, Amsterdam was a highly urbanised metropolis, attracting many immigrants from the rural areas of Holland and from other provinces of the Netherlands, as well as from abroad, mainly from German-speaking regions. Since both the geographical and the demographic situations may influence the language of the letters sent to and from Amsterdam, this justifies distinguishing the two separate categories and allows us to establish possible differences.

A careful selection from the letters transcribed for the two periods guaranteed an appropriate, although, due to the material available, not equal, representation of both male and female writers, of various social classes and of various age groups, which will become evident in the case studies of the following chapters. Compiling our corpus and labelling the letter writers according to social variables including region, we not only encountered the problems mentioned above, we also had to solve the crucial issue of the autograph or non-autograph status of letters.
6. The autograph status of letters, corpora and presentation of results

One of the familiar problems in historical-sociolinguistic research is the uncertainty about the autograph status of texts, which, for instance, Hernández-Campoy & Schilling (2012: 68–69) discuss for the Paston Letters, a collection of Middle-English private letters. The same problem applies to later periods and other language areas. In the case of our Dutch seventeenth- and eighteenth-century private letters, too, we should not hastily match specific language use with the sender of a letter, but take into account the possible literacy or illiteracy of the sender. Compiling our digital corpus thus involved complex research into the autograph or non-autograph status of the letters.

Although the rate of literacy in the Northern Netherlands was high compared to other European countries at the time, part of the population could neither read nor
write.\textsuperscript{13} We have to bear in mind also that those who were able to read may not have had any writing skills, as reading and writing were taught in succession, not simultaneously (Blaak 2009: 3–4; Kuijpers 1997: 501). On the basis of signature studies of marriage contracts, it is commonly estimated that two-thirds of the male population and one-third of the female population in the Northern Netherlands were able to write in the second half of the seventeenth century (Frijhoff & Spies 1999: 237). Literacy was also socially stratified, in that one-third of the lower ranks and two-thirds of the higher ranks of society were literate in the later part of the seventeenth century (Frijhoff & Spies 1999: 238). By around 1800, literacy had increased to about 80 per cent of the male and 60 per cent of the female population (Kloek & Mijnhardt 2001: 18). These literates, however, may not all have used their reading and writing skills regularly and thus differed in reading and writing experience. Literacy, therefore, should be considered as a continuum of various degrees, varying from almost illiterate, limited reading skills, limited writing skills to advanced reading and writing skills. For writing a letter, both illiterates and those with insufficient writing experience had to rely on more experienced writers, professional writers or what we call 'social writers', relatives and friends with more advanced writing skills. In sum, as part of the seventeenth- and eighteenth-century population was either illiterate or partly literate or had insufficient writing experience, we had to establish first of all whether or not the letters were written by the senders themselves.

In order to deal with this autograph problem, we developed the so-called Leiden Identification Procedure (LIP) (Nobels & van der Wal 2012). This procedure, which combines script and content analysis, was applied to our whole corpus of seventeenth-century letters. We will give a few examples to illustrate our method. An autograph or non-autograph is easily established whenever a sender of a letter refers to the writing process (‘my bad handwriting is due to darkness in the time of night’, for instance), or explicitly states that the letter is self-written, or that the letter was written by order of someone else. A non-autograph can also be determined whenever a clumsily written signature closes a letter in a fluent and professional handwriting. Most cases, however, are more complex and require detailed investigation, for instance whenever we encountered letters by the same sender, written in different handwritings, or whenever we found letters in identical handwritings, sent by different individuals.\textsuperscript{14} In these cases, and in fact in all other cases of single letters in which no direct content or form clue indicates its autograph or non-autograph status, we had to find external information.

\textsuperscript{13} Around 1800, the Northern Netherlands, Scandinavia, Iceland, Prussia and Scotland had fewer than 30 per cent male illiterates, a much lower percentage than, for example, the Southern Netherlands, England, Ireland and France (cf. Boonstra 1993: 22–23; see also Graff 1987: 173–248).

\textsuperscript{14} Identical handwritings were traced by making use of the GIWIS tool (Groningen Intelligent Writer Identification System) that was developed for our research programme by Axel Brink, member of the artificial intelligent team directed by Lambert Schomaker (University of Groningen).
about the senders in order to classify the letters. Official documents such as wills or marriage contracts may provide us with samples of the sender’s signature and handwriting which allow for comparison with the letters in order to draw conclusions about their autograph or non-autograph status. In cases in which we had insufficient proof, letters were ultimately labelled as having an ‘unclear status’.

We applied the LIP procedure to our seventeenth-century corpus, which resulted in three categories of letters: autographs, non-autographs and letters of unclear status. The category of autographs appeared to comprise writers from all four social classes, both men and women of different ages, thus allowing for historical-sociolinguistic research. Non-autographs were written by mainly unknown professional or social writers who were called upon by people with limited writing skills to write letters for them.15 For the eighteenth-century letters, the increasing rate of literacy makes the identification problem easier to solve. We found a high number of eighteenth-century autographs, although we still came across non-autograph letters written by professional or social writers (cf. van der Wal, Rutten & Simons 2012: 144–146). These non-autograph letters were excluded from our final digital corpus that, for the eighteenth century, comprises only autographs. Distinguishing autographs from non-autographs (and letters of unclear status) was thus an important methodological issue of our research.

The different subcorpora compiled are suitable for different types of research. The seventeenth-century autograph subcorpus and the whole eighteenth-century corpus are appropriate for research into social variation and will be used in the following

15. In a few cases, we were able to establish the identity of the social or professional writer. For illustrative examples see Chapter 5, Section 4 and van der Wal & Rutten (2013a).
Letters as Loot

chapters. The whole seventeenth-century and the eighteenth-century corpora can be used for research into regional variation, as the social and professional writers of non-autographs and of letters of unclear status were most likely to be found in the immediate environment of the sender. In almost all cases, these letters can, therefore, be labelled as related to the same region as the sender.\footnote{This connection could safely be made for non-autograph letters or letters of unclear status sent from the Netherlands, but not for such letters sent from abroad. In the latter case, the person that was called upon to write a letter, e.g. the ship’s writer, may not have originated from the same region as the sender of the letter. In that case the letters received the regional code ‘Unknown’ (cf. Nobels 2013: 31–33).} In passing, we note that the size of the corpora actually used in our various case studies also depends on the phenomenon investigated: detailed research on sounds and orthographical phenomena, for instance, was conducted on a limited subcorpus, as was research into clause chaining.\footnote{On the size of corpora, cf. Nurmi (2002) who also stresses that for lexical studies, for instance, much larger corpora are needed than for other phenomena.} A detailed account of the corpora used will be given in the various chapters. The basic numbers of the seventeenth- and eighteenth-century subcorpora are presented in Table 2; the data for the seventeenth-century autograph corpus, which is part of the seventeenth-century total subcorpus, are presented separately between brackets.

The subcorpora used for the present monograph were compiled from our corpus of almost 2,000 transcribed letters, following particular guidelines that restricted the number of words of individual writers.\footnote{Only a limited number of words was allowed for each individual writer, in order to avoid imbalance by particularly prolific writers (seventeenth-century writers max. 2,000 words, cf. Nobels 2013: 50; eighteenth-century writers max. 2,500 words, cf. Simons 2013: 86).} These subcorpora, the digital photographs and the related metadata are available at http://brievenalsbuit.inl.nl, an elaborate internet application with advanced search tools, developed in close collaboration with the Leiden Institute of Dutch Lexicology (INL). Being familiar with the impressive corpora available for the history of English, in particular the similar Corpora of Early English Correspondence (CEEC, CEECE, CEECSU), the reader may, at first sight, question the size of our subcorpora. On closer inspection, however, and taking into consideration the differences between the English and Dutch corpora, we can confidently present our letter corpora for historical-sociolinguistic research. Whereas the

<table>
<thead>
<tr>
<th>Period</th>
<th>Subcorpus</th>
<th>N letters</th>
<th>N words</th>
<th>N writers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1660s-1670s</td>
<td>Total</td>
<td>549</td>
<td>228,000</td>
<td>424</td>
</tr>
<tr>
<td>(1660s-1670s</td>
<td>Autograph</td>
<td>260</td>
<td>118,000</td>
<td>202</td>
</tr>
<tr>
<td>1770s-1780s</td>
<td>Total</td>
<td>384</td>
<td>196,500</td>
<td>292</td>
</tr>
<tr>
<td>Sum of both subcorpora</td>
<td>933</td>
<td>424,500</td>
<td>716</td>
<td></td>
</tr>
</tbody>
</table>
English corpora cover a time span of four centuries (1402–1800) with 5.3 million words, 1,066 writers and 11,819 letters (Palander-Collin et al. 2009: 14), our corpora cover two decades (that is one-twentieth of the time span of the English corpora) with approximately 424,000 words, 494 individual writers of autographs (or 716 writers including the professional or social writers) and 933 private letters, all of which have been transcribed from the manuscripts. A comparison with Elspaß’s German corpus points in the same direction, since this corpus of nineteenth-century letters covers a whole century with 700,000 words, 273 writers and 648 letters (Elspaß 2012b: 47).

Throughout the book, we will present our results in tables, cross-tables and figures, limiting the use of statistical tests to cases where the results seem to be close.

7. Outlook

Having discussed the principal aims of our research and the major issues of the present monograph, we would like to point out that individual chapters will also address additional topics related to the specific research phenomenon under discussion. Politeness and politeness theory, for instance, are a major theme in our chapter on forms of address (Chapter 6), and grammaticalisation figures in both the chapter on clause chaining (Chapter 7) and that on relative clauses (Chapter 8). Furthermore, although we concentrate on external factors of social and regional variation, internal factors such as clause type, phonological context and epistolary context will also be taken into consideration, for instance when examining relative clauses, apocope of final schwa and negation. We will also discuss the issue of contemporary literacy, for instance when dealing with clause chaining and relative clauses, and take into account the development of literacy which is assumed to have had a considerable impact on language histories. In fact, literacy, education and writing experience frequently constitute the background of our historical-sociolinguistic research, for instance when testing our assumption that educated writers use fewer typically oral elements in their written language. Finally, when presenting our results, we will, wherever appropriate, refer to parallels in other languages and compare our findings with previous work on English and German in particular (cf. publications by Austin (e.g. 1973a, 1973b), Fairman (e.g. 2007, 2012) and Elspaß (e.g. 2005, 2012b)).

The sociolinguistic variation of the past is generally difficult to discern, invisible in printed texts that have come down to us. This lack of insight into linguistic variants and variation is often assumed to be compensated for by analysing private letters written by people from various social ranks – letters that are supposed to reflect the everyday, more informal language of the time. The rediscovery of the confiscated documents from the past allowed us to put this hypothesis to the test by exploring a large selection of the private letters. The results of our challenging exploration will be presented in the present monograph in which we aim at both mapping variation and change and giving explanations for the variation and change revealed, taking into
consideration regional factors, gender, social class, age, diachronic change and epistolary context. As mentioned above, when achieving these results, our two ultimate goals were to fill the gaps in the history of Dutch and to make a contribution to the international historical-sociolinguistic debate. The exploration of the private letters as a complement to the history of Dutch and as a Dutch contribution to the international historical-sociolinguistic debate will be assessed and evaluated in the final chapter (Chapter 11).
CHAPTER 2

Sounds and signs
From local to supralocal usage

1. Introduction

In Chapter 1, we have seen that the letters contain traces of the spoken language of the time, phonological and morphological as well as syntactic. These traces, which do not appear in most of the published contemporary works, provide a window on the spoken language used in everyday communication. As such, the letters are of inestimable value. More central to our sociolinguistic concerns, however, is the question of the extent to which these traces of the spoken language occur. Do the letters represent the spoken language in a fairly straightforward and consistent way, or do we also find prototypically written language features? Can the language be characterised as primarily representing local, dialectal usage, or do they also contain elements of possibly supralocal writing systems? What is, in other words, the letters’ degree of orality (Rutten & van der Wal 2011)?

In Sections 2 and 3, we will explain why the degree of orality is an important research topic, and why we want to be able to distinguish between local and supralocal usage. In Sections 4 and 5, we report on a series of case studies concerning the interrelation of local sounds and forms, and (possibly) supralocal signs. In Section 4, we focus on letters to and from Zeeland, and in Section 5 on letters to and from North Holland, including Amsterdam. Section 6 consolidates the conclusions of the Sections 4 and 5.

In the present chapter, we are primarily interested in regional aspects of language use. Most of the research was carried out in an early phase of the project, when information on the regional origin of the letters was available, but detailed social information on the letter writers was still lacking. In a later stage, when social variables were assigned to the letters in the corpus, we carried out exploratory case studies into social variation, which will be reported on in Sections 5.2 and 5.3.

2. Standard languages, language standards and the degree of orality

Traditional language histories were often “primarily concerned with unification and standardisation processes”, and mainly on the basis of literary texts, as Elspaß (2007: 3)
Letters as Loot among others has stressed. Consequently, the language of large parts of the population remained unknown to us. To remedy this situation, the so-called language history from below (see Chapter 1, Section 2), focuses on non-literary everyday language, presumably found in ego-documents such as letters and diaries from lower- and middle-class writers. The language of these letters and diaries is considered to be as close to spoken language as possible (e.g. Elspaß 2005; Elspaß et al. 2007) and is assumed to give us a view on contemporary variation and change. From the perspective of the language history from below, the historical-sociolinguistic description and explanation of this variation and change will lead to an “alternative history” of the language (Watts & Trudgill 2002), an alternative for “the inexorable march towards a uniform standard” (Elspaß 2007: 3). Thus, this and our historical-sociolinguistic research is not driven by the “ideology of the standard language” (Milroy 2007), but aims to complement traditional linguistic history using new sources.

Whereas we may argue that an interpretation of variation in the past as a deviation from the emerging standard language should be avoided, we should not neglect the obvious presence of language standards in the past (Joseph 1987). Without the presupposition or the back projection of an emergent national standard, “consensus norms of regional speech communities” (Milroy 1994: 20) can be traced back, and “processes of supralocalization” (Nevalainen & Raumolin-Brunberg 2003: 161) have in fact formed the core of historical dialectology and historical sociolinguistics. The convergence of writing practices has been established even for medieval English (e.g. Nevalainen & Tieken-Boon van Ostade 2006). Similarly for Dutch, although charters by thirteenth- and fourteenth-century scribes can often be localised on the basis of certain dialect features, most of the scribes used variants from different dialect areas, and appear to have been acquainted with various regional spelling systems (Rem 2003: 261–263). Therefore, the beginning of a Dutch supraregional writing tradition is often dated as early as the thirteenth and fourteenth centuries (van der Sijs 2004: 554–556).

Similarly, studying the history of Late Middle and Early Modern English, Nevalainen & Raumolin-Brunberg (2003: 182, 183) claim that, although “supralocalizing features do not have any single path of transmission”, “most of the supralocalizing features followed the lead of the capital city”. In other words, the consequences of “Vertikalisierung” (von Polenz 1994: 136) could not be excluded. A remarkable case study in this respect is Austin (1994), who studied the eighteenth-century Clift letters, written by members of “a family of the artisan class, living in the remote south-western county of Cornwall” (Austin 1994: 285). Originally expecting to encounter “a language similar to speech or as near speech as we could now reasonably hope to find”, Austin concluded that this was not the case, rather, “as soon as these people took up a pen they framed their minds to a formal mode of thinking”, the result being that “their letters are neither markedly dialectal nor colloquial” (Austin 1994: 285, 286).

One of the main problems in historical sociolinguistics then, is to determine the degree of orality of documents from below, to clear up “the contrast between regional dialects and nonlocalizable language” (Nevalainen & Raumolin-Brunberg 2003: 161),
to disentangle “actual usage on the one hand and the taught written language on the other” (van der Wal 2007: 95), and “to assess the nature of the recording process in all possible and relevant ways and to evaluate and take into account its likely impact on the relationship between the speech event and the record, to reconstruct the speech event itself, as accurately as possible” (Schneider 2002: 68). Moreover, the internal development of writing systems, proceeding more or less independently from spoken language, should also be taken into account (Primus & Neef 2004: 133–134). To put it clearly: on the one hand, we should avoid teleological interpretations in terms of standardisation, while on the other hand, we should not unreflectively consider written data as spoken language transcripts, and be aware of possible processes of supralocalisation. Hence, the topic of the present chapter is: to what extent are data that we assume to be as close to the spoken language as possible, dialectic? And to what extent are they supraregional? What, in other words, is their degree of orality?

3. The degree of orality in the history of Dutch

The degree of orality is a key research topic in the history of phonology and orthography that has lead to a number of leading publications. Seminal studies at the crossroads of historical phonology and historical graphematics have been carried out by Elmentaler (2000, 2001, 2003). Studying written language from the fourteenth to the seventeenth centuries from the Meuse/Rhine area and thereby focusing on professional town clerks, Elmentaler (2003) was able to explain many of the variational patterns he encountered by referring to phonology, drawing on insights from historical as well as present-day phonology and dialectology. Coexisting writing systems were found for the fifteenth century (Elmentaler 2000), each reflecting the spoken language to a considerable extent. Fischer (2000), however, found reflections of spoken language (Mundartreflexe) in sixteenth-century documents from Soest in North Rhine-Westphalia, but not systematically, and mainly in texts by non-professional writers.1 Regarding the writing systems involved, Elmentaler argued against an interpretation of orthography in terms of “Etymologisierung” and “Gleichschreibung der Morpheme” (2001: 310–311). At the end of the period he studied, however, and especially in the seventeenth century, there was a trend toward “Morphologisierung” (Elmentaler 2003: 286). Voeste (2007: 91), dating this trend somewhat earlier, claimed that “[d]uring the 16th century, German spelling underwent essential changes: the syllabic and the morphological spelling principles were established”. The rise of etymological, morphological and syllabic spelling principles implies a growing distance from phonological and phonetic writing systems. Von Polenz (2000: 159) already interpreted the late medieval

period as one in which Schreiblandschaften were indicative. These Schreiblandschaften entailed supraregional levelling (Ausgleich, variant reduction), and developed “weitgehend unabhängig” of the spoken language. In the Dutch normative tradition, which began in the sixteenth century, the phonetic principle first predominated, but was complemented by morphological spelling principles from the first fully-fledged grammar of Dutch, dating from 1584, onward (Dibbets 1985: 360–362). Similarly, Salmon (1999) described English orthography from the late fifteenth to the late eighteenth centuries in terms of a gradual development towards a standard spelling system, in which traces of spoken language are increasingly absent. In terms of this development, manuscript spelling lagged behind in comparison to spelling in printed books. Finally, awareness of standardised norms became even stronger in more recent times, also among less-privileged people (cf. Austin 1994; McColl Millar 2012).

Focusing on Dutch private letters from the seventeenth and eighteenth centuries, we should, considering the time period, reckon with supraregional levelling as well as with morphological, etymological and syllabical spelling principles. After all, the seventeenth century is traditionally considered to be a stage of the standardisation process in which linguistic variation in published books strongly decreased, while the eighteenth century would mark the definite breakthrough of a more or less uniform written code (e.g. de Vooy 1952: 129). On the other hand, as most of our letters were written by non-professional writers from all social ranks, reflections of spoken language may very well be found, and it may even be that much of the orthographical variation can be explained phonologically.

In recent research of sixteenth- and seventeenth-century Dutch on the basis of ego-documents, the orthography of ego-documents is assumed to reflect the spoken language rather straightforwardly. Drawing on theories of dialect contact and koineisation (Trudgill 1986, Kerswill 2002), and using letters and diaries from immigrant families and their offspring, studies such as Boyce & Howell (1996), Boyce Hendriks & Howell (2000), Goss & Howell (2006), and Howell (2006) tried to reconstruct the formation of new Dutch urban vernaculars in the western part of the Netherlands, mainly in the province of Holland (Amsterdam, The Hague). As the sixteenth and seventeenth centuries are characterised by mass immigration to Holland by people from other dialect and language areas, dialect levelling and koineisation obviously took place. In order to trace the ongoing levelling and focusing of variants, ego-documents were used, that is written language, while dialect levelling, focusing, and koineisation are spoken language features. To justify the use of these texts for examining spoken language features, they were assumed to be written “in pure Holland dialect” (Boyce Hendriks & Howell 2000: 273) and to be “vernacular letters” (Howell 2006: 219). In other words, the relation between phonology and orthography, and between phonemes and graphemes, was not so much the object of study as presupposed to be

---

2. The first fully-fledged grammar of Dutch was the Twe-spraack vande Nederduitsche letterkunst ‘Dialogue on the grammar of Dutch’, published anonymously in 1584; see Dibbets (1985).
Chapter 2. Sounds and signs

23

isomorphic. It was admitted, however, that “certain orthographic traditions originating in the southern Netherlands continued to influence the written language in Holland” (Howell 2006: 210). We would like to stress that exactly these orthographic traditions may veil dialectal distinctions, rendering phonological research on the basis of written documents problematic. Similar ideas on the problematic relationship between phonemes and graphemes are expressed by i.a. Milroy with regard to Middle English (1992: 161–167).

It is against this background of the diverging opinions on the respective influence of local dialects and supralocal writing systems that we will carry out case studies of various regional features in order to establish the degree of orality of our letters.

4. Case studies Zeeland

In this section, we report on case studies on letters to and from Zeeland. We selected stable dialect features, that is, features which in historical linguistic and dialectological research are commonly assumed to be stable features of the dialects in question, in the present as well as in the past. The features will be explained in the relevant sections. In 4.1, we introduce the subcorpora used. In Sections 4.2 to 4.6, we discuss two prominent features of Zeeland dialects. Section 4.7 sums up the conclusions.

4.1 The subcorpora used

For detailed research on sounds and orthographical phenomena, we deliberately selected a limited subcorpus: 99 letters to and from Zeeland from the seventeenth century and 28 letters from the eighteenth century. We will refer to these as the seventeenth-century Zeeland subcorpus and the eighteenth-century Zeeland subcorpus, respectively. See Table 1 for the composition of both subcorpora. Note that the proportion of letters linked to Zeeland is much smaller in the eighteenth century than in the seventeenth century, not just in these subcorpora, but in the letter collection in general. Note, furthermore, that when compiling the subcorpora we have in each case avoided an imbalance in the number of letters relative to the number of letter writers, so as to avoid a situation where one or a few writers would be overrepresented, and thus distort the data. In the present study, this means that the 28 eighteenth-century letters linked to Zeeland were produced by no less than 26 different writers. Thus, with all due reservations, these letters may provide us with a clear view of eighteenth-century Zeeland.

For our fine-grained analysis we started with the exploration of a similar corpus for each period: 29 letters for the seventeenth century and 28 letters for the eighteenth century (see Sections 4.3 and 4.4). The subsequent research on all 99 seventeenth-century letters linked to the Zeeland region (at the time of research, early 2010; Section 4.2) did not reveal any striking differences.
Table 1. The Zeeland subcorpora.

<table>
<thead>
<tr>
<th></th>
<th>N letters</th>
<th>N words</th>
<th>Av N words/letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th-c. Zeeland subcorpus</td>
<td>99</td>
<td>45,000</td>
<td>455</td>
</tr>
<tr>
<td>18th-c. Zeeland subcorpus</td>
<td>28</td>
<td>13,000</td>
<td>464</td>
</tr>
</tbody>
</table>

epistolary practices. The right-hand column in Table 1, showing the average number of words per letter, is intended merely to demonstrate the comparability of both subcorpora, which allows us to compare tokens per letter for the two subcorpora.

The letters were selected solely for their link with Zeeland, considering them as representative of the everyday written language of Zeeland. Since most of these letters were written by merchants, and therefore by upper middle class males, research on social variables such as class and gender is not possible. It is feasible, however, to carry out research into regional variation, especially since most letter writers contribute only one letter, thereby increasing the subcorpora’s representativity. Our question about the extent to which oral features occurred in written language can be properly investigated, not least because the letters contain many instances of phonologically interesting features which can be straightforwardly linked to local usage.

The 99 letters from the seventeenth-century Zeeland subcorpus were used for case studies on *h*-dropping (Section 4.2) and the representation of etymologically different long *e*’s (Section 4.5). For a more fine-grained analysis of *h*-dropping, we used a subsection of the seventeenth-century Zeeland subcorpus, consisting of 29 letters (Section 4.3). The 28 letters making up the eighteenth-century Zeeland subcorpus were used for both the case study on *h*-dropping (Section 4.4) and the case study on the long *e*’s (Section 4.6).

4.2 *H*-dropping in the seventeenth century: First case study

In this section, we will introduce the feature of *h*-dropping and present the results of the first case study on the 99 letters from the seventeenth-century Zeeland subcorpus (Table 1). This study was carried out in order to gain a general overview of the prominence of *h*-dropping in letters to and from Zeeland.

Like many dialects from the south and the south-west of the Dutch language area, the Zeeland dialects are characterised by the absence of the phoneme *h*, i.e. the glottal fricative, as in *emel* instead of *hemel* “heaven” (cf. de Wulf et al. 2005: maps 214, 215). The phenomenon of *h*-dropping is also well-known from English historical phonology and dialectology (Lass 1992: 61–63). Middle Dutch texts, especially from the south-west of the language area (Flanders), often display deletion and prothesis of *h*. Schönfeld/Van Loey (1954: 93) note *hic* “I”, instead of *ic*, which already appeared in eleventh-century Old West Flemish. As will become clear from the examples below and from the case studies, Middle Dutch writing practices, where *h*-deletion and *h*-prothesis are commonly found, appear to have been continued into the Early Modern
period. It is significant, in this respect, that characters from the south were sometimes stereotyped by *h*-dropping in seventeenth- and eighteenth-century literature (see Rutten 2008 for some examples and discussion).

In the letters to and from Zeeland, four types of orthographical effects were found that were related to the non-phonemic status of *h* in the writer’s dialect:

- First and foremost: prevocalic deletion of ⟨h⟩ as in *andt* instead of *handt* “hand”, *adde* instead of *hadde* “had”, *uswrouwe* for *huswrouwe* “housewife”, and *eel* for *heel* “whole”. Prevocalic deletion of ⟨h⟩ in orthography points to deletion of [h] in pronunciation, so these instances are clear cases of dialect spelling.

- The second type is prothesis of ⟨h⟩ before vowels as in *hacht* instead of *acht* “eight”, *houde* for *oude* “old”, and *huere* for *ueren* “hours”. Although there is some lexically diffuse variation in present-day dialects (de Wulf et al. 2005: map 216; Weijnen 1966: 248), in that a prevocalic ‘hypercorrect’ [h] sometimes appears in spoken language, prothesis of ⟨h⟩ is also, and probably mainly, an orthographical phenomenon. In other words, prothesis of ⟨h⟩ indicates the influence not only of local or dialectal phonology, but also of supraregional orthography.

- The third and fourth types of orthographical effects related to non-phonemic *h* are the substitution of ⟨h⟩ for ⟨a⟩ as in *hpril* instead of *april* “April”, *hl* instead of *al* “all”, and the substitution of ⟨a⟩ for ⟨h⟩ as in *aebben* instead of *hebben* “have”, and *aoe* instead of *hoe* “how”. In the Dutch alphabet, ⟨a⟩ is pronounced [aː], and ⟨h⟩ is pronounced as [haː], but in the case of an *h*-less dialect, ⟨h⟩ is obviously pronounced [aː] as well, with two possible orthographical results representing the same sound. In these two cases of substitution, writers did not write their local dialect, but exhibit certain orthographical effects due to the learning of the alphabet on the basis of dialectal *h*-less phonology.

The first case study presented here concerns the presence of these four orthographical effects in the seventeenth-century Zeeland subcorpus. At least one token of one of the four possible orthographical effects was found in 63 out of 99 letters (64%). In total, the 63 letters produced 243 tokens, amounting to an average of 3.9 tokens per letter. Table 2 details the results per orthographical effect. Since the majority of letters contain at least one token, we may safely say that *h*-dropping must have been not just a prominent feature of the spoken language, but also that it occurs to a far greater extent in these private letters than in published texts of the time period. Middle Dutch writing practices obviously continued to exist well into the Early Modern period, whereas the language in published texts reveals supralocal variant reduction through the virtual absence of localisable features such as these.

---

4. It has been argued that more northern areas, such as Amsterdam, have also experienced a period when *h* was often deleted (Verdenius 1943). This controversial claim (de Tollenaere 1957) will be confirmed by our results in Section 5.2.4.
Letters as Loot

Table 2. Orthographical effects of non-phonemic h in seventeenth-century Zeeland.

<table>
<thead>
<tr>
<th></th>
<th>Deletion of ⟨h⟩</th>
<th>Prothesis of ⟨h⟩</th>
<th>Substitution of ⟨h⟩ for ⟨a⟩</th>
<th>Substitution of ⟨a⟩ for ⟨h⟩</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N letters with effect</td>
<td>46 (46%)</td>
<td>34 (34%)</td>
<td>8 (8%)</td>
<td>2 (2%)</td>
<td>63 (64%)</td>
</tr>
<tr>
<td>N tokens (total)</td>
<td>133 (55%)</td>
<td>75 (31%)</td>
<td>26 (11%)</td>
<td>10 (4%)</td>
<td>243 (100%)</td>
</tr>
<tr>
<td>N tokens (per letter)</td>
<td>2.9</td>
<td>2.2</td>
<td>3.3</td>
<td>5</td>
<td>3.9</td>
</tr>
</tbody>
</table>

As can be seen in Table 2, the two most prominent effects are deletion of ⟨h⟩, which occurs in 46 letters (46% of 99 letters), and prothesis of ⟨h⟩, which features in 34 letters (34%). The other two effects are relatively marginal, appearing in less than 10% of the 99 letters. Substitution of ⟨a⟩ for ⟨h⟩ was found in only two letters. The prominence of the first two effects is also clear from the token count, where deletion and prothesis make up 55% and 31% of the total number of 243 tokens, respectively. Whereas almost two-thirds (64%) of the Zeeland letters contain at least one orthographical effect of h-dropping, none of the four possible effects appears in the majority of the letters. Furthermore, the average token count points to only a few instances per letter. In the next section, we will develop a more detailed picture of the phenomenon of h-dropping, in order to find out whether we are actually dealing with a quite prominent feature (which has an effect in 64% of the letters), or, on the contrary, with what eventually proves to be a fairly marginal feature after all (with only a handful of occurrences per letter).

4.3 H-dropping in the seventeenth century: Second case study

Having obtained a general overview of the prominence of orthographical effects related to h-dropping, we present the results of a more fine-grained analysis that was conducted on a subsection of 29 letters of the seventeenth-century Zeeland subcorpus. We found at least one of the four possible orthographical effects in 20 of these 29 letters. This amounts to 69% of the letters, which is in accordance with the 64% as established in Section 4.2 (Table 2). In the 20 letters with at least one orthographical effect, we mainly found cases of h-deletion and h-prothesis, along with some instances of substitution of ⟨h⟩ for ⟨a⟩. These will be discussed in detail in the following sections (4.3.1, 4.3.2, 4.3.3). In this subsection of 29 letters, no tokens were found of substitution of ⟨a⟩ for ⟨h⟩, which was in any event a fairly marginal phenomenon (cf. Table 2).

4.3.1 Deletion of h

Of the 29 letters, 14 contain one or more tokens of deletion of prevocalic ⟨h⟩, which amounts to 48%. In the 99 letters of the seventeenth-century Zeeland subcorpus,
deletion occurred in 46% of the letters (Table 2). Usually, and again in accordance with the results in Table 2, letters contain two or three tokens of deletion.

In subsequent analyses, we will focus on these 14 letters, since they show variation of both prevocalic deletion (ø+V) and supraregionally written forms with initial ⟨h⟩ (⟨h⟩+V) and therefore give insight into the extent to which dialect spellings occur in the letters. We extracted all words which are written with initial ⟨h⟩ in the supraregional code, distinguishing between two different realisations, viz. ø+V and ⟨h⟩+V. In Table 3, we present the full results of the 14 letters with on the left side the instances of ø+V with their frequencies, and on the right side the instances of ⟨h⟩+V with their frequencies.

Table 3. Instances of ø+V and of ⟨h⟩+V in 14 letters from Zeeland.

<table>
<thead>
<tr>
<th>Letter</th>
<th>Examples</th>
<th>ø+V N</th>
<th>Examples</th>
<th>⟨h⟩+V N</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ant, aest</td>
<td>2</td>
<td>husvrouwe (3), husvrou, verhopende, het (12), hebbe (4), herten, hooren (2), hij (3), hooger, ghehad (2), heel, heef, hoogh, hoe (2), hebben, hooghtsen, hope (2)</td>
<td>38</td>
<td>40</td>
</tr>
<tr>
<td>2</td>
<td>geadt, adde (2), alf</td>
<td>4</td>
<td>alhier, hebt, het (3), hij (3), hier (3), heeft (2), haelen, huijs, gehaelt, hebe, alderhoochsten</td>
<td>18</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>ope (3), oope (2), adde (3), alf, aes</td>
<td>10</td>
<td>alhier, hier (2), hebben (4), hebbe (5), het (2), hoe (2), heeft (2), handt, heere, here, hij (2), husvrouwe</td>
<td>24</td>
<td>34</td>
</tr>
<tr>
<td>4</td>
<td>terandt, andt, geadt (2), adde (2), anden, veraelt, ope, verope</td>
<td>10</td>
<td>hij, hier, hebbe (4), hebben (2), heeft (2), gesontheijt (3), alhier, hier (2), het (6), hoe, boeckhouder, houijsvrouw, hoogst</td>
<td>26</td>
<td>36</td>
</tr>
<tr>
<td>5</td>
<td>eel, andt, adde</td>
<td>3</td>
<td>hebben, verhoop, het (8), heerten (2), hooren, heft, heef, hed, hoe (2), hoe, hy, hopen, gesonhet (2), hebbe, hope, heer (2), hier</td>
<td>29</td>
<td>32</td>
</tr>
<tr>
<td>6</td>
<td>uswrouwe, alf (2)</td>
<td>3</td>
<td>harten (2), gesontheit (4), heel (2), hij (2), verhoepen, het (6), heeft (3), hadde, hadt, hoepe (4), hemel, hoe, heere(n) (3), droefheit, hebt, hout, hier, houwen, behouden, hondert (2), haer (6)</td>
<td>45</td>
<td>48</td>
</tr>
<tr>
<td>7</td>
<td>oppe</td>
<td>1</td>
<td>hebben (2), verhopn, het (3), hertelijk, hooren, herten, gehoerd, hadde</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>eel</td>
<td>1</td>
<td>het (4), huis, heere(n) (4), gesonheijt, verhoope, verhooren, herte(n) (2), hooren, hoope, hebben, heeft, hier (2), verhuist, hadde, heijt, hollant, hondert, huisvrouwe</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td>Letter</td>
<td>Examples</td>
<td>$\varnothing$+V N</td>
<td>Examples</td>
<td>$\langle h\rangle$+V N</td>
<td>Total</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
<td>------------------</td>
<td>----------</td>
<td>------------------</td>
<td>-------</td>
</tr>
<tr>
<td>9</td>
<td>eer</td>
<td>1</td>
<td>gesontheijt, droevheij, heeft (3), heb(b)en (2), hoepende, heeren, hadden, heidenen (2), sterffhuis, hebt, handen, hooeren, bewinhebber, het, herte, droeffheijt, hvisvrou, huisvrou</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>10</td>
<td>ellevoedt</td>
<td>1</td>
<td>huisvrouwe (2), heeren, gesontheijt, verhoope, verhooren, hooren, bockhouwer, hier (2), hebbe, het (2), hij (2), heijt (3), hast, hondert, heij, heere</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>11</td>
<td>adde, eel, oogehvij, ollant, ondert</td>
<td>5</td>
<td>hvysvrouwe, verhoopen (2), hoope (2), het (12), herten, hooren, hebben (5), hebbe, haest (4), haer (3), hijer (2), hebbe, hvij, hvys (2), heft, handen, hem (2), heeft (3), hastighe, hoe, hy, behovde, hvysvrouwe</td>
<td>50</td>
<td>55</td>
</tr>
<tr>
<td>12</td>
<td>verope</td>
<td>1</td>
<td>het (5), hooren (2), hebben, gesontheijt (2), heeft (2), hebbe (3), hij (2), haest, behoude</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>13</td>
<td>alfven</td>
<td>1</td>
<td>herte (2), hoope (2), heere (2), gesontheijt, hebbe (2), gehadt, hij (2), heeft, haest, het (8), gelegentheijt, hebt (2), hout, houde, hier (2), gehoude, haer (2), behoude</td>
<td>33</td>
<td>34</td>
</tr>
<tr>
<td>14</td>
<td>eel</td>
<td>1</td>
<td>het (4), hoope (4), gesontheijt (2), herte, hoope, heere, verhoore, hebbe (5), hebt (2), gehadt, heeft (2), hij (3), hem (2), behoude, huijsvrouwe</td>
<td>31</td>
<td>32</td>
</tr>
<tr>
<td>N Total</td>
<td></td>
<td>44</td>
<td></td>
<td>394</td>
<td>438</td>
</tr>
<tr>
<td>% Total</td>
<td></td>
<td>10</td>
<td></td>
<td>90</td>
<td>100</td>
</tr>
</tbody>
</table>

In these 14 letters, no fewer than 44 instances of $\varnothing$+V were found where, supralocally, $\langle h\rangle$+V would be expected. This is a clear indication that zero was a normal option in the writers’ spoken language. Assuming that $h$-deletion was a stable dialect feature, as it is nowadays, and hence that any prevocalic $h$ could be deleted, we can compare the 44 instances to the 394 positions in the same 14 letters where the supralocal spelling $\langle h\rangle$+V appears. Our findings demonstrate that zero spellings, unambiguously revealing the spoken language, represent 10% of the phonologically possible deletions, that is, of words with a vowel which is preceded by [h] in other varieties of Dutch, and which is therefore preceded by $\langle h\rangle$ in supralocal writing systems. Put differently, letter writers who deleted $\langle h\rangle$ occasionally, in accordance with their local dialect, still
produced prevocalic initial ⟨h⟩ in 90% of the cases where it could have been deleted, and where presumably [h] was in fact not produced in the spoken language.

Turning to the individual letters, it is striking that in most cases there are only one or a few instances of ø+V. Moreover, only two letters (numbers 3 and 4) are responsible for 20 of the 44 tokens. Both letters contain 10 tokens, and in both letters, these still represent a minority compared to the 24 and 26 supralocal realisations, respectively. The first of these two letters is written by Cornelis Kien, whose letter is full of other Zeeland elements such as unrounded stieren for sturen “send”, unrounded diere for duur “expensive”, palatalised deur for door “through”, undiphthongised brandewin for brandewijn “brandy” (with [i] in the final syllable instead of [ei]), undiphthongised huisvrouwe for huisvrouwe “housewife” (with [y] instead of [œy] in the first syllable), and the original palatal vowel in geren “gladly” (with something like [e] instead of velarised gaarne with [a]). This letter writer incorporates quite a number of localisable variants, thereby not so strongly conforming with supraregional practice. In the other letter, written by Jan Dimmesz., we find fewer prototypically Zeeland features; one feature that we do find is undiphthongised sukr for suiker “sugar” (with [y] instead of [œy] in the first syllable). It should be noted that these features were (and are) also part of surrounding dialect areas, both to the north in Holland and to the south in Flanders, so that they cannot be considered uniquely Zeelandic. They diverge, however, from supralocal practice.

The results strongly suggest that we are dealing with a quite prominent dialect feature. At the same time, we observe that more than half of the writers (15 out of 29) always realise prevocalic ⟨h⟩, and those who do not do so consistently, still realise prevocalic ⟨h⟩ in a vast majority of all positions where deletion is possible. It can therefore be seen that the dialect feature of ⟨h⟩-dropping is far from systematically represented in the written language. From this we infer that Zeeland writers probably did not intend to put their Zeeland variety onto paper, but converged to supralocal practice instead, in which they very often succeeded.

Furthermore, there seems to be a specific phonological context triggering prevocalic ⟨h⟩-deletion, in that 37 out of 44 cases of ⟨h⟩-deletion (84%) in Table 3 co-occur with back vowels such as short [ɑ] as in alf “half”, long [a] as in aest “hurry”, short [ɔ] as in ollant “Holland”, and long [o] as in ope “hope”. Only 7 tokens co-occur with front vowels, e.g. eel “whole” with long [e], and uswrouwe “housewife” with long [y]. This suggests that back vowels promote ⟨h⟩-deletion more strongly than front vowels.

This is also suggested when investigating the proportion of back vowels that co-occur with ⟨h⟩-deletion. In the 14 letters under discussion, in total 108 V positions are taken by the back vowels short and long ɑ, and short and long o. Measuring the 37 instances of ⟨h⟩-deletion co-occurring with a back vowel against this total of 108 back vowels, we conclude that 34% of the back vowels co-occur with ⟨h⟩-deletion. This is a much higher proportion than the overall score of 10% prevocalic ⟨h⟩-deletion.

Note that the vowel spelled ⟨ou⟩, which frequently occurs in Table 3 and which probably represents a back vowel such as [ʌu] or [œu], is not taken into account.
Furthermore, there is one context where $h$-deletion prevails over $h$-realisation. No less than 22 of the 44 tokens of $h$-deletion co-occur not just with a back vowel, but with short [a] in particular (e.g. *ant* “hand”, *alf* “half”, *adde* “had”). Short [a] is the vowel produced farthest back in the mouth and phonetically closest to glottal [h]. Therefore, the difference between realisation with or without [h] must have been least audible in the spoken language when the following vowel was [a]. As can be seen in Table 3, the 14 letters under discussion contain 36 words where the V position is taken by [a]. It follows that 22 out of these 36 positions or 62% co-occur with $h$-deletion. This is the one single context where the local dialect orthographically prevails over supralocal writing practice.

4.3.2 Prothesis of $h$

Of the 29 letters used for a more fine-grained analysis, 11 contain one or more tokens of prothesis of ⟨h⟩. This amounts to 38%, resembling the 34% of letters containing $h$-prothesis in the Zeeland seventeenth-century subcorpus in general (Table 2). In accordance with the data in Table 2, these letters contain 1 to 3 tokens of $h$-prothesis.

As in Section 4.3.1, we will focus on these 11 letters in subsequent analyses, since these show variation of ø+V and ⟨h⟩+V, and therefore give insight into the extent to which dialect spellings occur in the letters. In Table 4, we present the results of these 11 letters: on the left-hand side the instances of localisable ⟨h⟩+V with their frequencies, and on the right-hand side the instances of ø+V with their frequencies. Building on the insight that the specific phonetic context influences these orthographical effects (Section 4.3.1), the localisable prothetic cases on the left-hand side were only compared to words beginning with a similar vowel. For instance, if a writer used *hacht* for *acht* “eight”, this case of prothetic ⟨h⟩ was only compared to words beginning with [a]. If a letter writer produced *hueren* for *ueren* “hours”, pronounced something like [yrə], we also looked at *uijt* “out”, since this was most probably pronounced something like [yt], without diphthongisation as in many seventeenth-century Brabantic and Hollandic dialects (e.g. [œyt]). We think we do most justice to the letter writers by restricting ourselves to similar phonological contexts. From the prothetic ⟨h⟩ which they produce, we only infer that they produce prothetic ⟨h⟩’s in such contexts, not that any initial vowel would be eligible.

Five letters with one or more tokens in Table 4 also appeared in Table 3: this applies to letters 1, 2, 6, 7, 8 and 11, corresponding to letters 1, 2, 6, 9, 11 and 14 in Table 3, respectively. In the 11 letters under discussion, we found 17 tokens of prothesis of ⟨h⟩. Of these, 12 co-occur with a back vowel such as [a], [ɔ], [u], [ʌu] or [ɔu]. 6 tokens co-occur with short [a]: four times in *hacht* “eight”, twice in *handere* “other”. Numbers are low here, but the tendency seems to resemble what we found in the previous section: short [a] most strongly promotes prothesis of ⟨h⟩.
Table 4. Instances of ⟨h⟩+V and of ø+V in 11 letters from Zeeland.

<table>
<thead>
<tr>
<th>Letter</th>
<th>Examples</th>
<th>N</th>
<th>⟨h⟩+V Examples</th>
<th>N</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>hueren (3)</td>
<td>3</td>
<td>ul (11), uit (2), uijtterste, u</td>
<td>15</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>houde</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>houde (3)</td>
<td>3</td>
<td>–</td>
<td>–</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>hacht</td>
<td>1</td>
<td>all (2), anders (4), als (15), afloop, ancker (3), alle (2), afgedaen (2), al (2), andere (3), af (2), ackoor, allsoo, anijs, apsentije, alles</td>
<td>41</td>
<td>42</td>
</tr>
<tr>
<td>5</td>
<td>hacht</td>
<td>1</td>
<td>apelles, als (3), al (3), alle</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>6</td>
<td>hachten</td>
<td>1</td>
<td>appril, al (4), anders, almachtich, als (4), alles, alle (3), altijt (20)</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>hoch</td>
<td>1</td>
<td>ons (6), op, of</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>7</td>
<td>hoeffen [oeffenen]</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>hijngelsman</td>
<td>1</td>
<td>in (5), is (7), ijs (2), ijck (24) [ent (2), en (18), ende] [21]</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td>9</td>
<td>handere (2)</td>
<td>2</td>
<td>Aprl, alle, als (2), altis, alsoo</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>10</td>
<td>her</td>
<td>1</td>
<td>en (8), engelsman (3)</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td>hacht</td>
<td>1</td>
<td>Aprel, alle, an, anders, al, als</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>N total</td>
<td>17</td>
<td></td>
<td>150</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td>% total</td>
<td>10</td>
<td></td>
<td>90</td>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>

Comparing the results for prothetic ⟨h⟩+V to supralocal ø+V in similar contexts, we conclude that 17 out of 167 phonologically possible positions are realised as ⟨h⟩+V. This is only 10%, similar to the 10% ⟨h⟩-deletion we found in Table 3. Limiting ourselves to short [a], we have 6 tokens of ⟨h⟩+⟨a⟩ as opposed to 78 tokens of ø+⟨a⟩ (see Table 4). This means that only 7% of all short [a]’s trigger prothesis of ⟨h⟩. Summing up, prothetic ⟨h⟩ is a fairly marginal phenomenon.

4.3.3 Substitution of ⟨h⟩ for ⟨a⟩

Substitution of ⟨h⟩ for ⟨a⟩ appears in only 3 of the 29 letters of the present subsection (10%). All instances are presented in Table 5. The first two letters both have one token

---

6. The letter writer produces hijngelsman “Englishman” with the vowel ⟨ij⟩, clearly suggesting the not uncommon pronunciation [I]; we therefore selected initial [I] for comparison, as in in “in”, is “is” and ijck “I”. Still, (h)engelsman with initial [e]/⟨e⟩ would also be possible, especially because it represents supralocal pronunciation and orthography. Accordingly, we added words with initial ⟨e⟩ in square brackets, which we did not incorporate in the overall calculation.
Table 5. Instances of substitution of ⟨h⟩ for ⟨a⟩ in 3 letters from Zeeland.

<table>
<thead>
<tr>
<th>Letter</th>
<th>Examples</th>
<th>⟨h⟩ for ⟨a⟩ N</th>
<th>⟨a⟩ (similar context) Examples</th>
<th>⟨a⟩ N</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Hprils</td>
<td>1</td>
<td>alle, anders, almachtich, alln, als (5)</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>Hl</td>
<td>1</td>
<td>april, anders</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>ontfhngen, versthén (2), hls (2), lhet, mhlkander, vrlhng</td>
<td>8</td>
<td>any ⟨a⟩</td>
<td>109</td>
<td>117</td>
</tr>
<tr>
<td>N Total</td>
<td></td>
<td>10</td>
<td></td>
<td>120</td>
<td>130</td>
</tr>
<tr>
<td>% Total</td>
<td></td>
<td>8</td>
<td></td>
<td>92</td>
<td>100</td>
</tr>
</tbody>
</table>

in initial position. We therefore compared these tokens to other words with initial ⟨a⟩. The third letter has tokens of ⟨h⟩ for ⟨a⟩ both in initial and in medial position, so we compared the tokens to any ⟨a⟩ appearing in that letter.

The ten tokens attested on the left-hand side of Table 5 represent 8% of all ⟨a⟩ positions in similar contexts. Naturally, as the numbers are low, we should treat them with caution here. The writer of the third letter is responsible for most tokens. Even if it appears to be the case that he may have randomly substituted ⟨h⟩ for any ⟨a⟩, he still produced 109 instances of ⟨a⟩ in line with supralocal practice. Summing up, substitution of ⟨h⟩ for ⟨a⟩ is a marginal phenomenon, even more so than ⟨h⟩-prothesis.

4.4 H-dropping in the eighteenth century, and diachronically

In this section, we discuss the orthographical effects of ⟨h⟩-dropping in the 28 letters which make up the eighteenth-century Zeeland subcorpus. We will also compare the results to the seventeenth-century results from the previous section so as to develop a diachronic perspective on the orthographical effects of ⟨h⟩-dropping in Zeeland.

The attested orthographical effects are the same as in the seventeenth century: deletion and prothesis of ⟨h⟩, and substitution of ⟨h⟩ for ⟨a⟩ and vice versa. We did a case study similar to the one reported in Section 4.2 above, so as to gain a better understanding of the prominence of these effects in the eighteenth century. We present the results in Table 6, giving the number and proportion of letters with a particular orthographical effect, as well as the token count and the average token count per letter.

As in the seventeenth-century Zeeland subcorpus (Table 2, Section 4.2), the respective orthographical effects occur in less than half of all the letters. Note that, in total, only 39% of the letters contain at least one effect. This was 64% in the seventeenth century. Similar to the seventeenth-century results, deletion and prothesis make up the greater part of the attested affects, but substitution of ⟨h⟩ for ⟨a⟩, which made up only 11% of the tokens in the seventeenth century, is now responsible for 30%. We will come back to this below.
Table 6. Orthographical effects of non-phonemic h in eighteenth-century Zeeland.

<table>
<thead>
<tr>
<th></th>
<th>Deletion of ⟨h⟩</th>
<th>Prothesis of ⟨h⟩</th>
<th>Substitution of ⟨h⟩ for ⟨a⟩</th>
<th>Substitution of ⟨a⟩ for ⟨h⟩</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N letters with effect</td>
<td>10 (36%)</td>
<td>4 (14%)</td>
<td>3 (11%)</td>
<td>2 (7%)</td>
<td>11 (39%)</td>
</tr>
<tr>
<td>N tokens (total)</td>
<td>32 (35%)</td>
<td>24 (26%)</td>
<td>27 (30%)</td>
<td>8 (9%)</td>
<td>91 (100%)</td>
</tr>
<tr>
<td>N tokens (per letter)</td>
<td>3.2</td>
<td>6</td>
<td>9</td>
<td>4</td>
<td>8.3</td>
</tr>
</tbody>
</table>

Figure 1. Orthographical effects of non-phonemic h in seventeenth- and eighteenth-century Zeeland.

Firstly, we look at the diachronic differences. In Figure 1, we present the proportion of letters with at least one orthographical effect (Total), and for every separate effect, distributed per century. The number of letters with at least one orthographical effect linked to the non-phonemic status of h sharply drops in the eighteenth century. Whereas 64% of the seventeenth-century letters contained at least one such effect, this applies to 39% of the eighteenth-century letters. Also, numbers for the two most important features, h-deletion and h-prothesis, drop from 46% to 36%, and from 34% to 14%, respectively. The overall picture is one of a decrease in local, and an increase in supralocal spellings.

The decrease in local spellings, and the concomitant convergence to supralocal practice, become clear from Figure 1. As to the increase in local spellings in the last two categories, ⟨h⟩ for ⟨a⟩ and ⟨a⟩ for ⟨h⟩, it should be noted that there are still only
Letters as Loot

a few letters in the eighteenth century that exhibit these characteristics (cf. Table 6). Where numbers are so low, a supposedly great increase is easily suggested.

What also becomes clear from Table 6, when compared to Table 2 in Section 4.2 above, is the rise in the average token frequency per letter, in particular with respect to \(h\)-prothesis and substitution of \(\langle h \rangle\) for \(\langle a \rangle\). In the seventeenth century, the average token frequency was 2.2 and 3.3, respectively. In the eighteenth century, these have risen to 6 and 9, respectively. Consequently, the average token frequency for all features has risen from 3.9 in the seventeenth century to 8.3 in the eighteenth century. This does not imply, however, a general increase in localisable spellings. The alleged increase appears to be solely due to two letters which are responsible for an extremely large number of spellings related to non-phonemic \(h\). The data for these letters are given in Table 7.

Table 7. Data for two letters from 18th-century Zeeland.

<table>
<thead>
<tr>
<th>Letter</th>
<th>Deletion of (\langle h \rangle)</th>
<th>Prothesis of (\langle h \rangle)</th>
<th>Substitution of (\langle h \rangle) for (\langle a \rangle)</th>
<th>Substitution of (\langle a \rangle) for (\langle h \rangle)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14</td>
<td>7</td>
<td>16</td>
<td>–</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>13</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Sub total</td>
<td>15</td>
<td>20</td>
<td>26</td>
<td>2</td>
</tr>
<tr>
<td>Total 18th-c. subcorpus</td>
<td>32</td>
<td>24</td>
<td>27</td>
<td>8</td>
</tr>
</tbody>
</table>

As is clear from Table 7, the writers of these two letters are responsible for almost all tokens of \(h\)-prothesis (20/24) as well as for almost all tokens of substitution of \(\langle h \rangle\) for \(\langle a \rangle\) (26/27). These letters do not contain an astonishing number of other Zeeland characteristics. The first one, written by “Helena Sap nu Adriaansen”, does contain unrounded \(dier\) for \(duur\) “expensive”, undiphthongised \(ter\ wile\) for \(terwijd\) “while”, and palatal \(merte\) for \(maart\) “March”. In the other letter, written by Daniel de Wolff, the most striking feature not in line with supraregional practices is the substitution of \(\langle e \rangle\) for \(\langle i \rangle\) as in \(begenen\) for \(beginnen\) “begin”, \(hen\) voor \(in\) “in”, and \(Meddelburg\) for \(Middelburg\). In both letters the most outstanding localisable features are without any doubt the orthographical effects of non-phonemic \(h\). The alleged rise in local spellings, when counting the average token frequency per letter, should be seen against the background of these two letters.

Finally, the fact that only a few letters are responsible for a relatively large number of features was also clear from Table 3 with the seventeenth-century results (Section 4.3.1). This is a theme we will return to in Sections 5.2.1, 5.2.2, 5.2.3, 5.2.4 and 5.7 below.

4.5 Long e’s in the seventeenth century

In the present section and in Section 4.6, we will discuss another salient characteristic of Zeeland dialects and its orthographical representation in the seventeenth- and
eighteenth-century Zeeland subcorpus: the feature of etymologically distinct long e’s. First, we will explain this feature in more detail before discussing the seventeenth-century results. In Section 4.6 we will deal with the eighteenth-century results and develop the diachronic perspective.

Many Dutch dialects, the southern ones in particular, maintain the phonological difference between lengthened ê out of originally short vowels in open syllables, and é out of the Wgm. diphthong *ai. In Dutch historical linguistics, the lengthened é is traditionally called softlong, while ê is called sharplong. We will retain to these terms.

- **Softlong** é developed through lengthening of the short vowels [ε] and [I]. Examples are the vowels in the first syllables of leven “live”, rekenen “count”, and hemel “heaven”. Compare German leben, rechnen, himmel.

- **Sharplong** ê out of Wgm. diphthong *ai is found in e.g. steen “stone”, een “one”, heten “be called”. Compare German Stein, ein, heißer.

In present-day standard Dutch, these phonemes have merged into one and the same long [e:]. The merger dates back at least to the end of the sixteenth century, and probably started in Amsterdam (Rutten 2009a). Dialects from Zeeland distinguish the two long e’s even to the present day, along the lines of most dialects that distinguish both e’s, with softlong é being rather monophthongal (e.g. [e:]), and sharplong ê being a diphthong (e.g. [Iə], cf. van Bree 1987: 103–104). According to Goossens et al. (2000, map 21), softlong week “week” has a homorganic diphthong [ei], and sharplong steen “stone” (map 128) has a centring diphthong [eə].

Contrary to the feature of h-dropping (Sections 4.2, 4.3, 4.4), where supraregional practice was very much in favour of dialects maintaining h, in the case of the long e, supraregional practice distinguished the two etymologically distinct long e’s. Southern writing practices from the sixteenth century onward distinguished both e’s in open syllables by doubling the sharplong ê in such cases.

- Thus, this **phonology-based** system had ⟨e⟩ for softlong ê in open syllables, for instance leven “live”, rekenen “count”, and hemel “heaven”, as opposed to ⟨ee⟩ for sharplong ê in open syllables steenen “stones”, een “one”, heeten “be called”.

This practice was also often maintained in the North, although new practices came into being, reflecting the ongoing merger (Rutten 2009a).

- A **morphologically**-oriented system came into use, in which open syllables received ⟨ee⟩ not on phonological grounds, but in analogy with their closed counterparts, e.g. sharplong steenen “stones” because of the singular steen “stone”, and also softlong leeven “live” because of the singular leef “live”, as opposed to softlong hemel “heaven”.

- In **syllabic** systems, the choice of the grapheme depends on the syllable structure. There are two basic options: either ⟨e⟩ or ⟨ee⟩ in open syllables, e.g. sharplong stenen as well as softlong leven and hemel, or sharplong steenen as well as softlong leven and heemel.
In what follows, we will first establish to what extent the phonological difference between softlong en sharplong e, which we assume to have been part of seventeenth-century Zeeland dialects, is represented in the seventeenth-century Zeeland subcorpus. We hypothesise that it is maintained in these letters, in line with the spoken language, and is moreover backed up by a phonology-based supraregional writing tradition which also distinguishes reflexes of the different phonemes. Consequently, we expect to find ⟨e⟩ for ̃e in open syllables, and ⟨ee⟩ for ̃e in open syllables. The phonology-based system will prove to be dominant when scrutinising the present subcorpus as a whole, but morphological and syllabic systems will also be shown to be in use when focusing on the writing system per individual letter.

We extracted all words with a long e in open syllable from the seventeenth-century Zeeland subcorpus of 99 letters, and then divided the data according to both the etymological origin of the vowel (either ̃e or ̃e) and the orthographical representation (either ⟨e⟩ or ⟨ee⟩). Table 8 gives the results.

The results for sharplong ̃e in open syllables are very convincing: nine out of ten are spelled ⟨ee⟩. For softlong ̃e in open syllables, numbers are less straightforward, yet two-thirds of all tokens have ⟨e⟩. The dialectic phonological distinction, founded on the etymological difference, is fairly well represented in the orthography, and is in line with supralocal, originally Southern practices.

It does not seem probable that there are synchronic phonological reasons for the orthographical variation in the representation of softlong ̃e. One could, for instance, hypothesise that the following consonant may have influenced the vowel, and that the actual distribution in the spoken language need not be entirely in line with the etymological origin of the vowels. Focusing on following consonants, however, no clear pattern arises. There are, for example, a large number of ⟨ee⟩ spellings for softlong ̃e followed by a dental, as in weeten “know” (30 tokens) and meede “with, also” (18 tokens). But these two lexemes are in any event among the most frequent, and are still found much more often with ⟨e⟩: weten appears 61 times in the letters, mede 90 times. Still,

Table 8. Frequency of ⟨e⟩ and ⟨ee⟩ for ̃e and ̃e in open syllables in seventeenth-century Zeeland.

<table>
<thead>
<tr>
<th></th>
<th>Softlong ̃e</th>
<th></th>
<th>Sharplong ̃e</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>⟨e⟩</td>
<td>715</td>
<td>67</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>⟨ee⟩</td>
<td>358</td>
<td>33</td>
<td>148</td>
<td>90</td>
</tr>
<tr>
<td>Total</td>
<td>1073</td>
<td>100</td>
<td>165</td>
<td>100</td>
</tr>
</tbody>
</table>

7. ⟨e⟩ and ⟨ee⟩ were by far the most frequent graphemes, though occasionally another grapheme was used, such as ⟨eij⟩ for ̃e. These were counted as instances of ⟨ee⟩ since they only stress the diphthongal realisation of reflexes of ̃e. See also below, Table 9 and the explanation there.
we have to make a fundamental reservation: there may have been phonological variation partly generating the orthographical variation, unknown to us and not reconstructable on the basis of the available data.

The general conclusion following from Table 8 is that the phonological difference between the long e’s is fairly well represented in the seventeenth-century Zeeland subcorpus, and hence that a phonology-based orthographical system predominates. This conclusion needs to be refined, however, when focusing on the writing system per letter. As explained above, we have to reckon not only with phonological spelling systems, but also with morphological and syllabic systems and, moreover, with a writing system which might be termed variable. After all, letter writers may use none of the afore-mentioned systems, and apparently randomly distribute ⟨e⟩ and ⟨ee⟩ over both long e’s instead. In such cases, we refer to the spelling system as variable.

When deciding on the writing system used in individual letters, we allowed for two exceptions. Accordingly, if a letter consistently contains ⟨ee⟩ for sharplong ê in open syllables, and ⟨e⟩ for softlong ê in open syllables, with two exceptions, for instance softlong weeten “know” instead of weten, and sharplong stenen “stones” instead of steenen, we will still call the system used in this letter phonology-based or phonological. The reason for this is that demanding no exceptions at all would imply that almost no letter contained any discernible spelling system. This would not only be an unfortunate route to take, it would also do little justice to a language stage which is characterised by much less uniformity and much less normativity than the period from about 1800 onward. At first, the number of maximally two exceptions was chosen arbitrarily; it could equally have been one or three. In practice, it proved to be quite useful, in that most letters clearly contained either the variable system, that is with much more than two unclear or irregular tokens, or one of the other systems, with indeed one or two exceptions. Note, finally, that we have only taken into consideration letters with examples from both categories, that is, from both softlong ê and sharplong ê in open syllables. If one of the categories was empty, the spelling system was left undecided. Out of the 99 letters in the seventeenth-century Zeeland subcorpus, 27 lacked sufficient forms, which left us with 72 letters with one of the spelling systems mentioned above.

Of these 72 letters, 31 (43%) have the phonology-based spelling system described above, with ⟨e⟩ for softlong ê in open syllable, and ⟨ee⟩ for sharplong ê in open syllable. Only 14 of these are completely regular, in that there were no exceptions at all. In Table 9, we give three examples of letters with the phonology-based system. Note that in the second letter, the grapheme ⟨eij⟩ appears in meijne “mean (verb)”. This grapheme rarely occurs, the occasions when it does appear being mainly in the words meijne, alleijn “alone”, and deijl “part”, and only for sharplong ê. The grapheme underlines the diphthongal realisation of the phoneme, and we deemed it natural to incorporate it in the results for ⟨ee⟩.
Table 9. Phonology-based distribution of ⟨e⟩ and ⟨ee⟩ in three letters from seventeenth-century Zeeland.

<table>
<thead>
<tr>
<th>Letter</th>
<th>Softlong ē</th>
<th>Sharplong ė</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ⟨e⟩</td>
<td>desen (2), weder (2), beter, wecke, afschepen, schepen (2), geschreven, weten, wegens, vrede, wesen (2), wese (2), mede, afbrecke, gelegentheyt, wetten [weten]</td>
<td>-</td>
</tr>
<tr>
<td>⟨ee⟩</td>
<td>-</td>
<td>alreede, verleen</td>
</tr>
<tr>
<td>2 ⟨e⟩</td>
<td>geve, mede, weten, geschreven, gelegentheijjt (2), even, weghe, beter, geseten, versekert, sekere, onsekere, versekeren, spreken, voornemen (2), vernedert, betere, bevelen, neven</td>
<td>-</td>
</tr>
<tr>
<td>⟨ee⟩</td>
<td>-</td>
<td>vreesen, meijne</td>
</tr>
<tr>
<td>3 ⟨e⟩</td>
<td>reden, negerin (2), weder (2), tegen (2), verdwenen, mede, wegens</td>
<td>rede</td>
</tr>
<tr>
<td>⟨ee⟩</td>
<td>Weegens</td>
<td>einege, seelant</td>
</tr>
</tbody>
</table>

The third letter in Table 9 exemplifies the phonological system with two exceptions allowed. In both categories, the letter has one form diverging from the main pattern: softlong weegens “because of” is spelled ⟨ee⟩, whereas sharplong rede “roadstead” is spelled ⟨e⟩. These exceptions are in italics in Table 9.

As well as the phonology-based systems, morphologisation of the spelling system can also be determined. Morphological systems are often combined with the phonological system in that softlong ē in open syllables is mostly spelled ⟨e⟩, and sharplong ē in open syllables mainly ⟨ee⟩, but softlong ē is rendered ⟨ee⟩ in case of analogy with morphologically simpler counterparts in closed syllables, such as the singular, the first person or the root. Of the 72 letters in the seventeenth-century Zeeland subcorpus with a determinable spelling system, there are 10 letters (14%) with such influence of morphology. In Table 10, we present the results of two sample letters where influence of morphology can be discerned.

Table 10. Morphological influences on the distribution of ⟨e⟩ and ⟨ee⟩ in two letters from seventeenth-century Zeeland.

<table>
<thead>
<tr>
<th>Letter</th>
<th>Softlong ē</th>
<th>Sharplong ē</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 ⟨e⟩</td>
<td>desember, rekeninge (2), seventis [sic], schepen weegend, weese (2), weesen, weeken, verteeren, heere, seeker</td>
<td>meenninge</td>
</tr>
<tr>
<td>⟨ee⟩</td>
<td>-</td>
<td>alleene, teekent</td>
</tr>
<tr>
<td>2 ⟨e⟩</td>
<td>ghelheijdtjdt, desen, werlet [wereld], toeghenegen leeuwen, weesen, heeten [eten], heere (2), seegent, eeuenwel</td>
<td>-</td>
</tr>
<tr>
<td>⟨ee⟩</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The third letter in Table 9 exemplifies the phonological system with two exceptions allowed. In both categories, the letter has one form diverging from the main pattern: softlong weegens “because of” is spelled ⟨ee⟩, whereas sharplong rede “roadstead” is spelled ⟨e⟩. These exceptions are in italics in Table 9.

As well as the phonology-based systems, morphologisation of the spelling system can also be determined. Morphological systems are often combined with the phonological system in that softlong ē in open syllables is mostly spelled ⟨e⟩, and sharplong ē in open syllables mainly ⟨ee⟩, but softlong ē is rendered ⟨ee⟩ in case of analogy with morphologically simpler counterparts in closed syllables, such as the singular, the first person or the root. Of the 72 letters in the seventeenth-century Zeeland subcorpus with a determinable spelling system, there are 10 letters (14%) with such influence of morphology. In Table 10, we present the results of two sample letters where influence of morphology can be discerned.
In both letters, the phonological distribution of ⟨e⟩ and ⟨ee⟩ is breached whenever soft-long ē in open syllable is spelled ⟨ee⟩ for morphological (analitical) reasons. In the first letter, we find

- softlong weeggend “weighing” by analogy with the root/first person singular weeg
- softlong weese and weesen “be, are” by analogy with the root wee
- softlong weeken “weeks” by analogy with the singular week
- softlong verteeren “digest, consume” by analogy with the root/first person singular verteert
- softlong heere “lord” by analogy with the singular heer.

In the second letter in Table 10, we find

- softlong leeuwen “live” by analogy with the root leev or the first person singular leef
- softlong weesen “be” by analogy with the root wee
- softlong heeten “eat” by analogy with the root/first person singular eet
- softlong heere “lord” by analogy with the singular heer.

There are also exceptions in both letters, which we rendered in italics: softlong seeker “sure” in the first letter, and softlong seegent “blesses” and eeuwenwel “however” in the second letter. In these cases, ⟨ee⟩ is used while there is no morphologically plausible analogy with a root or a singular form.

The influence of the syllable structure can also be determined in the spelling strategies employed by letter writers from Zeeland. We have 10 out of 72 letters (14%) with a syllabic writing system, that is with consistently either ⟨e⟩ or ⟨ee⟩, irrespective of the etymology of the long e in question. Table 11 gives the results of two letters. Exceptions are in italics.

In the first letter, the writer uses ⟨e⟩ for both soft-long and sharp-long e’s. There is one exception, viz. beeter “better”. In the second letter, the system is ⟨ee⟩ for soft-long as well as for sharp-long e. Again there is one exception: reken “measure (verb, 1 sing)”. Note that we have incorporated ⟨eij⟩ spellings (alleijn “alone”, meijnt “thinks”) into the

<table>
<thead>
<tr>
<th>Letter</th>
<th>Softlong ē</th>
<th>Sharplong ē</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>⟨e⟩</td>
<td>rede, leren</td>
</tr>
<tr>
<td></td>
<td>⟨ee⟩</td>
<td>Beeter</td>
</tr>
<tr>
<td>2</td>
<td>⟨e⟩</td>
<td>Reken</td>
</tr>
<tr>
<td></td>
<td>⟨ee⟩</td>
<td>deese (3), meede (5), eeten (2), verneemende, gesneede, gesteeken (2), neegen (2), seeuen, weederom, reeden, reegen (4), weese, weesen, weidersijts, neegers, beesich, beeter, tseedert, weete, weetende, heeren (2)</td>
</tr>
</tbody>
</table>
(ee) results as these underline the realisation of sharplong ê as a diphthong, as explained above. The (eij) spelling in particular shows that phonology is not entirely ruled out in these letters. The syllabic system of the second letter, however, mainly concerns the originally softlong ē, spelled ⟨ee⟩, not in line with the existing southern writing tradition.

Finally, 21 out of 72 letters (29%) with a determinable spelling system contain enough data in that they contain ample examples of both softlong and sharplong e, yet the spelling system applied does not appear to be phonology-based, nor does it exhibit influence of morphologisation or syllabification. These letters have the variable spelling system, which means that the writer randomly distributed ⟨e⟩ and ⟨ee⟩. The results of two example letters are given in Table 12. The conclusion that a writer freely varied in the choice of graphemes is not surprising in itself. From the history of orthography and phonology it is well-known that variation was a common phenomenon in the seventeenth century that was also found with consonants (e.g. ⟨g⟩ vs ⟨gh⟩, ⟨k⟩ vs ⟨ck⟩, ⟨kw⟩ vs ⟨qu⟩, ⟨s⟩ vs. ⟨z⟩, ⟨l⟩ vs ⟨v⟩; cf. Gledhill 1973). Note that, as in Tables 10 and 11, the variation mainly applies to softlong ē. Sharplong ê is spelled ⟨ee⟩ in 90% of all instances in the seventeenth-century Zeeland subcorpus (Table 8), and thus it is also written in the letters in Table 12. These letters provide examples of the variable distribution of ⟨e⟩ and ⟨ee⟩.

Whereas the sharplong vowels are all spelled ⟨ee⟩, variation is found in the case of the softlong ē’s. The first letter could be said to contain the phonological system, in line with the supraregional tradition, with three exceptions however, while we allow only two exceptions. More importantly, the letter contains bevelle “command (verb)” as well as beveele, which points to free variation. In the case of a morphological system, there would also be three exceptions: bevelle, which has a root beveel, and on the other hand overleede “died” and meede “with, also”, for which there is no plausible analogical form. In the case of a syllabic system, with consistently either ⟨ee⟩ or ⟨e⟩, there would of course again be three exceptions. Similarly, a syllabic system is not at work in the second letter, nor is there evidence of a phonological system. There would be far too many exceptions. A morphological system is also improbable, as it would need to be explained why geven “give” and wesen “be, are” are spelled ⟨e⟩, whereas there are root forms and/or first person singular forms with ⟨ee⟩. On the other hand, beeter “better” and meede “with, also” should have been written ⟨e⟩.

Table 12. Variable distribution of ⟨e⟩ and ⟨ee⟩ in two letters from seventeenth-century Zeeland.

<table>
<thead>
<tr>
<th>Letter</th>
<th>Softlong ē</th>
<th>Sharplong ê</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>⟨e⟩</td>
<td>dese, bevelle, werelt</td>
</tr>
<tr>
<td></td>
<td>⟨ee⟩</td>
<td>overleede, meede, beveele</td>
</tr>
<tr>
<td>2</td>
<td>⟨e⟩</td>
<td>desen, geschreven, wederom, geven, wesen (2)</td>
</tr>
<tr>
<td></td>
<td>⟨ee⟩</td>
<td>weeten (2), beeter, meede, heeren (2), heere (2), eenen, alreede, seelant</td>
</tr>
</tbody>
</table>
Table 13. Writing systems in 72 letters from seventeenth-century Zeeland (99 letters in total).

<table>
<thead>
<tr>
<th></th>
<th>Phonology-based</th>
<th>Syllabic</th>
<th>Morphological</th>
<th>Variable</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>〈e〉</td>
<td>〈ee〉</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N letters</td>
<td>31</td>
<td>3</td>
<td>7</td>
<td>10</td>
<td>21</td>
</tr>
<tr>
<td>(Proportion)</td>
<td>(43%)</td>
<td>(4%)</td>
<td>(10%)</td>
<td>(14%)</td>
<td>(14%)</td>
</tr>
</tbody>
</table>

Summing up the analysis of the long e’s so far, we have established spelling systems for 72 out of 99 letters making up the seventeenth-century Zeeland subcorpus; see Table 13. The other 27 letters did not contain enough forms to draw conclusions on the writing system used.

The spelling system most frequently used in these letters can be said to be phonology based (43%). The influence of morphological (10 letters, 14%) and syllabic systems (10 letters, 14%) was found in 20 letters in total (28%). On the one hand, phonology also played an important role in these letters in that sharplong è’s were often still written 〈ee〉, in accordance with the supraregional tradition. On the other hand, morphologisation and syllabification are unambiguously linked to and arise from the written code. They represent what could be termed the graphemisation of writing systems, that is, the reduction of phonological considerations and the increase in choices directly linked to the written code. In the next section, we will see that this graphemisation continued into the eighteenth century.

4.6 Long e’s in the eighteenth century, and diachronically

As we saw in Section 4.5, the phonology-based system predominated in the seventeenth-century Zeeland subcorpus, while the overall distribution of 〈e〉 and 〈ee〉 also favoured the importance of phonology. Moreover, the etymologically sound distribution of 〈e〉 and 〈ee〉 was largely in line with supraregional writing practices. At the same time, the influence of morphological and syllabic writing systems was determined. We will now examine the eighteenth-century Zeeland subcorpus and compare the outcome with the seventeenth-century results. First, we will present an overview of the eighteenth-century results, then we will focus more closely on the writing system per individual letter.

We intend to make comparisons with the seventeenth-century results throughout the present section. We therefore carried out a similar case study on the 28 letters of the eighteenth-century Zeeland subcorpus. First, we extracted all words with a long e in open syllable, and established the choice of grapheme for sharplong è and sharplong ê, respectively. Table 14 presents the results.
Table 14. Frequency of \( \langle e \rangle \) and \( \langle ee \rangle \) for \( \ddot{e} \) and \( \dot{e} \) in open syllables in eighteenth-century Zeeland.

<table>
<thead>
<tr>
<th></th>
<th>Softlong ( \dot{e} )</th>
<th></th>
<th>Sharplong ( \ddot{e} )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>( \langle e \rangle )</td>
<td>150</td>
<td>38</td>
<td>5</td>
</tr>
<tr>
<td>( \langle ee \rangle )</td>
<td>246</td>
<td>62</td>
<td>42</td>
</tr>
<tr>
<td>Total</td>
<td>396</td>
<td>100</td>
<td>47</td>
</tr>
</tbody>
</table>

As in the seventeenth-century (Table 8, Section 4.5), \( \langle ee \rangle \) is used for sharplong \( \ddot{e} \) in open syllables in about nine out of ten instances. As before (Section 4.5), we incorporated into \( \langle ee \rangle \) a few tokens where a slightly different grapheme underlines the diphthongal realisation of the sharplong \( 
\dot{e} \), such as teyken “sign” with \( \langle ey \rangle \) instead of \( \langle ee \rangle \).

For the softlong \( \dot{e} \), however, the pattern is reversed. Whereas the seventeenth-century letters contained \( \langle e \rangle \) in two-thirds of all instances, the eighteenth-century letters have \( \langle ee \rangle \) in about two-thirds of all instances. The diachronic difference also becomes clear from Figure 2, where we compare the choice of grapheme per century, based on Tables 8 and 14.

As stated, the pattern for sharplong \( \ddot{e} \) in open syllable is very stable, while the pattern for softlong \( \dot{e} \) changes dramatically. In the eighteenth century, both softlong and sharplong \( e \) are spelled \( \langle ee \rangle \) in most cases, which runs counter the phonological difference between the phonemes in most Zeeland dialects. It is also not in line with supra-regional practices originating from the southern language area, based on this very phonological difference. For an explanation of this remarkable difference, we turn to the writing system per individual letter.

Figure 2. Sharplong \( \ddot{e} \) and softlong \( \dot{e} \) in seventeenth- and eighteenth-century Zeeland.
Focusing on the writing system per letter, it emerges that the striking difference in the representation of soft long ē’s is due to a general shift towards more syllabic writing systems, in other words, to the ongoing graphemisation of the written language (see Section 4.5). Of the 28 letters making up the eighteenth-century Zeeland subcorpus, 25 provide enough data on long ē’s for establishing the writing system used. Of these 25 letters, no less than 12 (48%) contain a syllabic writing system. The main system is ⟨ee⟩ in open syllables, regardless of phonological (etymological) differences. In Table 15, we present the data of one such letter. Again, we allow for two exceptions to the writing system used, which we put in italics (see Section 4.5).

The writer of the letter in Table 15 opts for ⟨ee⟩ in open syllables for both long ē’s. He consistently applies this system, with two exceptions. One of the exceptions, dezes “this (genitive)” has counterparts in the category of ⟨ee⟩, where both deeze and deezen “this” appear.

We were able to determine the writing systems of 25 (out of 28) letters from the 18th-century Zeeland subcorpus, distinguishing between phonology-based systems, systems with clear morphological or syllabic influence, and variable systems. In Table 16, the writing systems of these 25 letters are presented, and compared to the seventeenth-century results, taken from Table 13 in Section 4.5.

Table 15. Syllabic influences on the distribution of ⟨e⟩ and ⟨ee⟩ in a letter from eighteenth-century Zeeland.

<table>
<thead>
<tr>
<th>Softlong ē</th>
<th>Sharplong ē</th>
</tr>
</thead>
<tbody>
<tr>
<td>⟨e⟩</td>
<td>edele, dezes</td>
</tr>
<tr>
<td>⟨ee⟩</td>
<td>geschreven (2), weezen (2), weeten, Scheepen, geweezene, Leeven (3), meede, beeter, reeden, deeze (2), deezen, leezen, heere, heeren</td>
</tr>
</tbody>
</table>

Table 16. Writing systems in 25 letters from eighteenth-century Zeeland (28 letters in total), compared to 72 letters from the seventeenth-century.

<table>
<thead>
<tr>
<th></th>
<th>Phonology-based</th>
<th>Syllabic</th>
<th>Morphological</th>
<th>Variable</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>⟨e⟩</td>
<td>⟨ee⟩</td>
<td>Total</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total syllabic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17th c.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N letters</td>
<td>31</td>
<td>3</td>
<td>7</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>(Proportion)</td>
<td>(43%)</td>
<td>(4%)</td>
<td>(10%)</td>
<td>(14%)</td>
<td>(14%)</td>
</tr>
<tr>
<td>18th c.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N letters</td>
<td>3</td>
<td>2</td>
<td>10</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>(Proportion)</td>
<td>(12%)</td>
<td>(8%)</td>
<td>(40%)</td>
<td>(48%)</td>
<td>(100%)</td>
</tr>
</tbody>
</table>
Apart from the large number of letters with variation as the apparent system, there is a striking difference between the seventeenth-century results and the eighteenth-century data. While the phonology-based system was in use in 43% of the seventeenth-century letters, it is only used in 12% of the eighteenth-century letters. At the same time, syllabic systems, and most prominently ⟨ee⟩ in open syllable, rise from 14% in the seventeenth century to 48% in the eighteenth century. Summing up, phonological spelling gives way to syllabic spelling. The written language graphemicises, in that a prototypical aspect of the written code, viz. orthography choices in open syllables, prevails against the phonological differences of the spoken code. Note that softlong and sharplong e’s are still distinguished in most Zeeland dialects. Therefore, the change in writing systems cannot be interpreted as the orthographical reflex of an ongoing change in the spoken language. The phonological difference being maintained in the spoken language, the syllabification of the writing system is merely graphematic. This implies that the written code is moving away from the spoken language, and that the extent to which the spoken language is represented in the written language decreases.

In Section 5.6, we will explain this superimposition (Überschichtung, cf. Elmentaler 2003) of phonology-based writing systems by syllabic strategies as a result of convergence to practices in North Holland and/or Amsterdam, where the etymologically different long e’s had in fact merged. In Section 5.6, therefore, we will return to the representation of long e’s, both in Amsterdam and in Zeeland. First, we sum up the main conclusions of Section 4 on the Zeeland subcorpora before proceeding with case studies on letters to and from North Holland.

### 4.7 Conclusions

One main conclusion following from the case study on h-dropping in letters linked to Zeeland is that many letters contain a handful of tokens linked to the absence of h as a phoneme, but only a few letters contain more than a handful of tokens. Moreover, even if there are more than a handful of tokens, the majority of historical-phonologically possible positions are still spelled in accordance with the supraregional practice. Consequently, the letters from Zeeland definitely contain a lot of interference from the spoken language, that is, from local dialects, but all in all writing systems seem to be outspokenly supralocal. Furthermore, there is a decrease in unambiguously local features from the seventeenth to the eighteenth century.

In this respect, the results of the study on the long e’s are very similar. We witnessed a decrease in phonology-based writing systems in favour of syllabification of the spelling. Again, spoken language features give way to prototypically written language features. The decrease in phonology-based writing systems came as a surprise, since there was in fact a supralocal writing tradition based upon the etymological difference between the two long e’s. In Section 5.6, we will show that the change from phonology-based to syllabic should be interpreted as a case of convergence to the
dialect of North Holland, in particular that of Amsterdam. Summing up, traces of spoken language notwithstanding, the writing systems employed in the Zeeland sub-corpora are characterised by convergence to supralocal practice.

By way of illustration, we round off this section with some self-corrections found in letters from the seventeenth-century Zeeland subcorpus. These self-corrections provide further evidence that supralocal practices rendered the language of the letters less local, thus reducing the letters’ degree of orality.

In one letter there is an instance of a horizontal strike-through ⟨h⟩ “h ick”, where perhaps the writer wanted to render ick “I” as hick, which would have been a token of h-prothesis. Alternatively, the writer may have begun hoepe “hope”, and then decided to put in the subject ick “I” first. See Figure 3 where the text reads:  h ick hoepe “I hope” and note the difference between the first letter of the line and the initial ⟨h⟩ of hoepe.

Figure 3. Fragment of a letter by Jan Jaspersen, dated 9 December 1664.

Another letter shows correction of the negative particle hen, with h-prothesis, into en. See Figure 4 with the vertical strike-through of the ⟨h⟩. The transcription and translation of the fragment are given in example (1).

(1) reviere van surijnaeme sijn gearrijveert, wij river of Surinam are arrived we <ken hebben tot nogh toe hier van geen viant neg have until yet to here from no enemy gehoort, soo dat wij teegenwoordigh hier heard so that we presently here ‘[we] have arrived [at] the river of Surinam. Until now, we have not heard about an enemy here. So that presently, we here...’

Figure 4. Fragment of a letter by Dirck Pottey, dated 11 September 1672.
Once again, in this case an alternative possibility presents itself: with the corrected 〈h〉 the writer may have falsely anticipated the verb hebben. Our third example, however, unambiguously shows a struck-through 〈h〉 which must reflect h-hesitation and could not be related to anticipation of a following word. The 〈h〉 is followed by a word with an initial vowel, viz. alder “most” (see Figure 5). The transcription and translation of the fragment are given in example (2).

Figure 5. Fragment of a letter by Gillis Block, dated 14 November 1664.

Figure 6. Fragment of a letter by Gillis Block, dated 14 November 1664.

(2) waert het aenders het waert mij van herten let
    were it different it were me from heart sorry
om te ooren dat wet ghij wel mij 〈h〉 alder[...]liijste maet
to to hear that know you well my 〈h〉 most[...]dearest mate
If it were different, I would be very sorry to hear that. This you know well my most dearest mate.

In the same letter, 〈h〉 is once followed by aebbe “have”, where 〈a〉 substitutes 〈h〉, contrary to supralocal practice, but the very fact that the writer corrects himself is proof of his awareness of the difference between local spoken language and supralocal written language (see Figure 6 with “Ick 〈h〉 aebbe”). Corrections such as these convincingly show us the clash of local dialects and supralocal writing practices.

5. Case studies North Holland and Amsterdam

In this section, we will examine dialect features in letters to and from North Holland, its capital Amsterdam included. As with the Zeeland letters discussed in Section 4, we selected stable dialect features, that is, features which in historical linguistics and dialectology are commonly assumed to be or to have been stable characteristics of the
North-Holland dialects, in the present as well as in the past. The features selected will be explained in the relevant sections. After having introduced the subcorpora used in Section 5.1, we report on a series of case studies in Sections 5.2 to 5.6. Section 5.7 sums up the conclusions.

5.1 The subcorpora used

It should be noted beforehand that we distinguish between Amsterdam, the main city in the province of North Holland, and the remaining part of North Holland (cf. Chapter 1, Section 5). As a metropolis, Amsterdam attracted a lot of immigrants, and it is therefore not self-evident that the language usage in Amsterdam still resembled the North-Holland dialects of the surrounding area. We compiled four subcorpora, which we termed the seventeenth-century North-Holland subcorpus, the seventeenth-century Amsterdam subcorpus, the eighteenth-century North-Holland subcorpus, and the eighteenth-century Amsterdam subcorpus. See Table 17 for the make-up of the four subcorpora.

Table 17. The North-Holland and Amsterdam subcorpora.

<table>
<thead>
<tr>
<th>Subcorpus</th>
<th>N letters</th>
<th>N words</th>
<th>Av N words/letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th-c. North-Holland subcorpus</td>
<td>39</td>
<td>18,500</td>
<td>474</td>
</tr>
<tr>
<td>17th-c. Amsterdam subcorpus</td>
<td>54</td>
<td>29,000</td>
<td>537</td>
</tr>
<tr>
<td>18th-c. North-Holland subcorpus</td>
<td>24</td>
<td>10,000</td>
<td>417</td>
</tr>
<tr>
<td>18th-c. Amsterdam subcorpus</td>
<td>166</td>
<td>80,000</td>
<td>482</td>
</tr>
</tbody>
</table>

The seventeenth-century North-Holland subcorpus contains letters sent to and from the towns of Alkmaar, Edam, Egmond aan Zee, Enkhuizen, Haarlem, Hoorn, Monnickendam and Nieuwewand. This subcorpus will form the heart of Sections 5.2, 5.3, 5.4 and 5.5 which deal with the \( \text{sk} \) feature, the \( \text{ft} \) feature, the \( \text{a} \)-like vowels and the prefix \( \text{ge} \) respectively. The eighteenth-century North-Holland subcorpus comprises letters sent to and from towns such as Alkmaar, Zijpe, Enkhuizen, Den Helder, Den Hoorn (on the island of Texel), Graft, De Rijp and Medemblik. It will be used in Sections 5.2 and 5.3.

The two Amsterdam subcorpora were used for subsequent case studies on the \( \text{sk} \) feature and the \( \text{ft} \) feature in Sections 5.2 and 5.3, as well as for the study of long \( e' \)’s in Section 5.6. Our study of the \( \text{a} \)-like vowels and the prefix \( \text{ge} \) in the seventeenth-century North-Holland subcorpus (Sections 5.4 and 5.5), resulted in so few tokens that we refrained from further examining these features in the eighteenth-century North-Holland subcorpus and in both Amsterdam subcorpora.

In Sections 5.2 and 5.3 we compare the results for North Holland with those for Amsterdam in order to find out whether there are any significant differences between Amsterdam and the North-Holland subcorpora; such differences may be due to
Amsterdam’s special status as a metropolitan city with a high degree of urbanisation. As the subcorpora comprise letters that can often be linked to social rank, we intended to explore them for social stratification and investigate possible effects of social class
and gender. Conducting the case studies, we found, however, such low frequencies of localisable features that systematic research on social class and gender effects was not feasible. Nevertheless, we will refer to social class and gender when discussing particular letters and individual writers.

As stated above, we used the Amsterdam subcorpora again for Section 5.6, where we focus on the long e’s which were also discussed in Sections 4.5 and 4.6 with regard to letters linked to Zeeland. Moreover, in Section 5.6 we will not use the North-Holland subcorpora, but solely the Amsterdam subcorpora, or part of these, since the merger of the two etymologically distinct long e’s is linked historically first and foremost to Amsterdam (Rutten 2009a). Therefore, Amsterdam provides the ultimate point of comparison with Zeeland, where the etymological difference between the two long e’s is maintained into the present. As the orthographical representation of the long e’s makes for fairly laborious case studies, we decided against investigating the full Amsterdam subcorpora. Instead, we randomly selected 50 letters to/from Amsterdam, viz. 30 from the seventeenth-century Amsterdam subcorpus, and 20 from the eighteenth-century Amsterdam subcorpus.

5.2 Germanic sk

In this section, we discuss the orthographical representation of reflexes of Gm. sk in the seventeenth-century North-Holland subcorpus (5.2.1), in the seventeenth-century Amsterdam subcorpus (5.2.2), in the eighteenth-century North-Holland subcorpus (5.2.3), and in the eighteenth-century Amsterdam subcorpus (5.2.4).

In the Middle Dutch period, the [sk] pronunciation for reflexes of Gm. sk was probably still widespread in the Dutch language area. The ⟨sc⟩ spelling was in any case commonly in use for sk in both initial and medial position. In final position, the ⟨sch⟩ spelling was very common, signaling an ongoing sound change. From the thirteenth century onward this spelling also turned up in initial and medial position (van Loey 1965: 116–117).

In many North-Holland dialects, Gm. sk is maintained as [sk] whereas most other varieties of Dutch as well as present-day standard Dutch have [sy] or [sx] in initial position, further reduced to [s] in medial and final position. In initial position, the postmedieval conventional spelling used to be and still is ⟨sch⟩, as in school “school”, pronounced as [syo:l] or [sxo:l], though with initial [sk] in North Holland. The conventional spelling used to be ⟨sch⟩ in medial and final position as well, but has been ⟨s⟩ since the mid-twentieth century: mensch “man”, and menschen “men” are now spelled mens and mensen, pronounced as [mens] and [mensa],8 respectively. Nowadays, most dialects in North Holland with initial [sk] have [s] or another ‘velarless’ sibilant in medial and final position (de Wulf et al. 2005: maps 9–14).

8. Final -n in verbs and plural nouns is deleted in many varieties of Dutch including present-day standard Dutch.
In Amsterdam, the [sk] pronunciation disappeared, yet Commandeur (1988: 163) notes that it still existed in the nineteenth century. It goes without saying that ⟨sch⟩ may also represent [sk], especially in dialect areas where [sk] is the normal pronunciation. When writers vary, however, in the use of ⟨sc⟩, ⟨sk⟩ and ⟨sch⟩, awareness of supralocal writing practices must be at work.

5.2.1 Seventeenth-century North Holland
In this first study, we extracted all instances of ⟨sc⟩, ⟨sk⟩ and ⟨sch⟩ from the 39 letters in the seventeenth-century North-Holland subcorpus, on the assumption that the unconventional signs ⟨sc⟩ and ⟨sk⟩ would reveal localisable sounds, while ⟨sch⟩ is the conventional, supralocal sign which does not provide unambiguous information on the pronunciation.

We found signs of [sk] pronunciation, such as bisscop “bishop”, vrienscap “friendship”, scrieft “writes”, wenske “wish (verb)”, scijp “ship”. In a dialect area that still maintains the [sk] pronunciation, such examples illustrate that writers may have put their local spoken language onto paper. At the same time, they show that Middle Dutch writing practices continued well into the Early Modern period. No more than 6/39 letters (15%) contain such signs of localisable [sk]. We present the results of these six letters in Table 18. We compare these results with all historical-phonologically possible positions where spelling such as ⟨sc⟩ or ⟨sk⟩ could have been used as a sign of [sk] pronunciation, but where the writer used supralocal ⟨sch⟩ instead.

<table>
<thead>
<tr>
<th>Letter</th>
<th>⟨sc⟩ or ⟨sk⟩</th>
<th>N</th>
<th>⟨sch⟩</th>
<th>N</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Bisscop</td>
<td>1</td>
<td>geschreven, winschinge, gansche (2), geschiet, schouwende, reeckenschap</td>
<td>7</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>2 Vrienscap</td>
<td>1</td>
<td>mensche (2), geschrieuien, schep, schien, Aenschien, wensche, wiensche</td>
<td>8</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>3 scrieft, wenske</td>
<td>2</td>
<td>schepen, schiep, schiep, Geschreuen, veruersschen, schap [schaap], mensche, schrieft, wiensche</td>
<td>9</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>4 kopscip, wecken [wensen]</td>
<td>2</td>
<td>schrijven (3), schrijft, bescheijdt, wensche, wenschen</td>
<td>7</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>5 scade, gassce [ganse], scrijve, scijnt, scrijft, gescildert</td>
<td>6</td>
<td>schp [schip]</td>
<td>1</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>6 scijp, gaescooten, scepen</td>
<td>3</td>
<td>–</td>
<td>–</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>N Total</td>
<td>15</td>
<td>32</td>
<td>47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Total</td>
<td>32</td>
<td>68</td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Spellings indicating [sk] pronunciation were found in six letters (15%). The other 33 letters (84%) only have the supralocal grapheme ⟨sch⟩. Whatever the pronunciation of the writers of these 33 letters may have been, though this was probably [sk], in the written language they kept to the supralocal sign ⟨sch⟩. They may have interpreted the grapheme ⟨sch⟩ as the representation of [sk], or they may have avoided ⟨sc⟩ and ⟨sk⟩ due to awareness of the localisability of these graphemes. This has to remain undecided. We will, therefore, focus on the six letters with graphemic variation, suggesting awareness of supralocal practice on the one hand, and interference of local dialects on the other hand. Three of these letters stem from Enkhuizen and three others from the town of Monnickendam. In letters 1, 2, 3 and 4 in Table 18, a minority of the possible positions is in fact realised with ⟨sc⟩ or ⟨sk⟩. In these letters, there are only one or two localisable tokens as opposed to seven, eight or nine instances of supralocal ⟨sch⟩. In letters 5 and 6, the opposite is the case. All in all, there are no more than 15 tokens revealing local pronunciation (32% of all possible positions), and 32 tokens with the supralocal sign (68%). Note that no fewer than six of the 15 tokens of ⟨sc⟩ and ⟨sk⟩ are produced by a single letter writer (letter 5). Summing up, the localisable spelling occurs in a minority of the letters, and within these letters in a minority of the historical-phonologically possible positions.

5.2.2 Seventeenth-century Amsterdam

An additional study was carried out on the seventeenth-century Amsterdam subcorpus which contains 54 letters by 47 writers from all social classes, men as well as women. The pronunciation [sk] can safely be assumed for seventeenth-century Amsterdam (Kloeke 1934–35: 7; Commandeur 1988: 163). Again, we extracted all words with initial ⟨sc⟩, ⟨sk⟩ and ⟨sch⟩. The overall results are presented in Table 19.

Table 19. Distribution of ⟨sc⟩ and ⟨sch⟩ in seventeenth-century Amsterdam.

<table>
<thead>
<tr>
<th></th>
<th>⟨sc⟩</th>
<th></th>
<th>⟨sch⟩</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>39</td>
<td>%</td>
<td>457</td>
<td>92</td>
<td>496</td>
<td>100</td>
</tr>
<tr>
<td>%</td>
<td>8</td>
<td></td>
<td>92</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The 54 letters of the seventeenth-century Amsterdam subcorpus contain 496 historical-phonological positions where local spellings such as ⟨sc⟩ and ⟨sk⟩ could have been used. Whereas ⟨sk⟩ does not appear at all, ⟨sc⟩ is used in 39 cases, amounting to 8% of all possible positions. Examples include *scrijft* “writes”, *gescreven* “written”, *orlogencep* “warship”, *afsceren* “tear off”. These 39 tokens are found in 4 out of 54 letters or 7%, which is even less than in the seventeenth-century North-Holland subcorpus, where 15% of the letters contained tokens of ⟨sc⟩ and/or ⟨sk⟩ (Section 5.2.1). Taking into account the individual tokens in each of these 4 letters, we present in Table 20 the frequencies of local ⟨sc⟩ and supralocal ⟨sch⟩.
As becomes clear from Table 20, the total of 39 tokens is largely produced by one letter writer, viz. Lijsbet Flipse Amelingh, the writer of the first letter in Table 20. She uses the grapheme \( \langle \text{sc} \rangle \) 31 times in her letter, which amounts to a majority of 67% of all possible positions. In the other three letters, \( \langle \text{sc} \rangle \) spellings are in the minority. Numbers are low here. Firstly, this confirms what the seventeenth-century North-Holland subcorpus (Section 5.2.1) already suggested, viz. that interference of localisable spoken language with the written code is the exception rather than the rule. Secondly, these low numbers render research into social variation difficult. Two of the letters in Table 20 were written by an LMC writer, two by a UMC writer. Note, however, that letters 1, 2 and 3 were written by women, and letter 4 by a man. In letter 4, the one token of \( \langle \text{sc} \rangle \) can reasonably be interpreted as a slip of the pen, since no fewer than 25 tokens of \( \langle \text{sch} \rangle \) are its counterpart. This means that in the seventeenth-century Amsterdam subcorpus, with all possible reservations, women appear to be more likely users of local spellings such as \( \langle \text{sc} \rangle \) than men, as \( \langle \text{sc} \rangle \) spellings are found in three (out of 24) letters by women, and in only one letter (out of 30) by a man. While this interpretation may seem somewhat tenuous considering the scarcity of data, it is in line with expectations about language awareness built on the social and gender distribution of writing experience; this will be explained in Chapter 4, and will in fact be a recurrent theme throughout the present book.

Table 20. Frequencies of \( \langle \text{sc} \rangle \) and \( \langle \text{sch} \rangle \) in four letters from seventeenth-century Amsterdam.

<table>
<thead>
<tr>
<th>Letter</th>
<th>( \langle \text{sc} \rangle )</th>
<th>( \langle \text{sch} \rangle )</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>1</td>
<td>31</td>
<td>67</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>42</td>
<td>7</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>33</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>43</td>
<td>51</td>
</tr>
</tbody>
</table>

5.2.3 Eighteenth-century North Holland

The results for the seventeenth-century North-Holland subcorpus showed that only six out of 39 letters (15%) contained localisable \( \langle \text{sc} \rangle \) and \( \langle \text{sk} \rangle \) spellings. In the eighteenth century, three out of 24 letters (13%) have instances of \( \langle \text{sc} \rangle \) (\( \langle \text{sk} \rangle \) does not appear). Table 21 presents the results for these letters.

The overall proportion of \( \langle \text{sc} \rangle \) spellings (83%) is fairly high, but this is due to two letters (numbers 2 and 3) by one writer, viz. Anna Visser. The first letter is written by Antje Cornelis, who contributes six fairly short letters to the eighteenth-century North-Holland subcorpus, and in only one letter varies between \( \langle \text{sch} \rangle \) and
Table 21. Instances of ⟨sc⟩ in three letters from North Holland.

<table>
<thead>
<tr>
<th>Letter</th>
<th>⟨sc⟩</th>
<th>N</th>
<th>⟨sch⟩</th>
<th>N</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>scoon, scrijft</td>
<td>2</td>
<td>schape, schraal</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>scip, gescoote, vrintscap, vrintscop, bescerming</td>
<td>5</td>
<td>–</td>
<td>–</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>scipper, sceeppe, frintscap</td>
<td>3</td>
<td>–</td>
<td>–</td>
<td>3</td>
</tr>
<tr>
<td>N Total</td>
<td></td>
<td>10</td>
<td></td>
<td>2</td>
<td>12</td>
</tr>
<tr>
<td>% Total</td>
<td></td>
<td>83</td>
<td></td>
<td>17</td>
<td>100</td>
</tr>
</tbody>
</table>

⟨sc⟩. In the other five letters, she uses ⟨sch⟩ 13 times. Summing up, ⟨sc⟩ seems to be a marginal phenomenon in eighteenth-century North Holland. This is corroborated by the overall results for the eighteenth-century North-Holland subcorpus, as presented in Table 22: ⟨sc⟩ is used in no more than 9% of all historical-phonologically possible positions.

Table 22. Distribution of ⟨sc⟩ and ⟨sch⟩ in eighteenth-century North Holland.

<table>
<thead>
<tr>
<th></th>
<th>⟨sc⟩</th>
<th>%</th>
<th>⟨sch⟩</th>
<th>%</th>
<th>Total</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>10</td>
<td>9</td>
<td>100</td>
<td>91</td>
<td>110</td>
<td>100</td>
</tr>
</tbody>
</table>

5.2.4 Eighteenth-century Amsterdam

A similar case study to that reported on in the previous section was carried out on the eighteenth-century Amsterdam subcorpus, since the pronunciation [sk] can also be assumed for eighteenth-century Amsterdam (Kloeke 1934–35: 7; Commandeur 1988: 163). Whereas this subcorpus with its 166 letters by 137 writers is much larger than the seventeenth-century Amsterdam subcorpus with its 54 letters, it contains only seven tokens of the localisable spellings ⟨sc⟩ and ⟨sk⟩. Examples include scip, skip “ship”, scryft “writes”, afscyt “farewell”. These seven tokens appear in four letters, that is, in 2% of all the letters in the subcorpus. Naturally, any social or gender interpretation would be groundless in this case. We can, however, conclude that the seventeenth-century minimal interference of local dialect with the written language even decreased in eighteenth-century Amsterdam. The overall results of the eighteenth-century Amsterdam subcorpus are given in Table 23. The eighteenth-century Amsterdam subcorpus contains 1213 instances of reflexes of Gm. sk in initial position. As can be seen in Table 23, 1206 instances, that is, 99% of all historical-phonologically possible positions are spelled ⟨sch⟩, in line with supralocal practice.
Table 23. Distribution of /sc//sk/ and /sch/ in eighteenth-century Amsterdam.

<table>
<thead>
<tr>
<th></th>
<th>/sc//sk/</th>
<th>/sch/</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>1206</td>
<td>99</td>
</tr>
<tr>
<td>1213</td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As stated before, the seven instances of /sc/ and /sk/ appear in four letters. Three letters contain only one token. There is one letter, however, with four tokens, viz. scryft “writes”, scryven “write”, afscyt “farewell”, and vrindschp “friendship”. Note that in vrindschp, ⟨h⟩ replaces ⟨a⟩ in the suffix. These four tokens even outnumber the two supralocal tokens in the letter, which are, remarkably, schryven “write” and afschyd “farewell”. The male writer of this letter, Zacharias Nyman, who was born in Amsterdam in 1747, is allocated to the LMC. We stress that Nyman’s letter is remarkable in at least two other respects. Firstly, it contains a second prominent dialect feature: h-dropping. Whereas h-dropping is usually associated with the south and the south-west of the language area, it has been suggested that Amsterdam also experienced an h-less period (Verdenius 1943, cf. de Tollenaere 1957). Nyman’s letter provides clear evidence that this was in fact the case. Of all seventeenth- and eighteenth-century letters from Amsterdam, this is the only letter with so many orthographical effects of h-dropping. Examples include aar instead of haar “her”, which is a case of h-deletion, and two instances of substitution of ⟨h⟩ for ⟨a⟩: hmpt for ampt “profession”, and vrindschp for vrindschap “friendship”. In three more examples, ⟨h⟩ substitutes ⟨a⟩ and simultaneously, prevocalic ⟨h⟩ is deleted: gehdt for gehadt “had (part.)”, hnden for handen “hands”, hndt for handt “hand”. The second remarkable characteristic of Nyman’s letter is the large number of epistolary formulae. As we will show in Chapter 4, the use of formulae strongly decreased from the seventeenth to the eighteenth century. In the eighteenth century, the use of formulae is largely restricted to letter writers with the least writing experience. Following the social and gender distribution of the writing experience, the formulae are mainly found in the LC and in the LMC, as well as among women. Summing up, Nyman appears to have been a less-experienced writer, less aware of supralocal practice, and thus regularly allowing dialect interference with his written language.

Finally, note that most tokens of /sc/ and /sk/ in the eighteenth-century subcorpus are concentrated in one letter, while most other letters contain no tokens, or only one. We have seen this pattern before and we will see it again. We will return to this in Section 6.

5.3 Germanic ft

The present section deals with the orthographical representation of reflexes of Gm. ft in the seventeenth-century North-Holland subcorpus (5.3.1), in the seventeenth-century Amsterdam subcorpus (5.3.2), in the eighteenth-century North-Holland subcorpus (5.3.3), and in the eighteenth-century Amsterdam subcorpus (5.3.4).
In most varieties of Dutch, including present-day standard Dutch, reflexes of Gm. \textit{ft} [ft] are rendered \textit{cht} [χt]. The change from \textit{ft} to \textit{cht} is one of the most prominent features distinguishing Dutch dialects from most German, English and Frisian dialects, although it also occurs in many western Low German and Central Franconian dialects. Compare

- Dutch \textit{kracht} “power”, \textit{lucht} “air”, \textit{zacht} “soft”, \textit{achter} “behind” with
- English \textit{craft}, \textit{soft}, \textit{after},
- German \textit{Kraft}, \textit{Luft}, \textit{sanft},
- and Frisian \textit{kreft}, \textit{loft}, \textit{seft}, \textit{efter}.

This change, probably dating back to the tenth century, did not reach Holland early on (Schönfeld/van Loey 1954: 96). In the present day, \textit{ch(t)} prevails in \textit{zacht} and \textit{achter} in Dutch dialects (de Wulf et al. 2005: maps 130, 131), but this differed in the past. Berte-loot (1983: maps 132, 133) shows that in the thirteenth century, \textit{koft} instead of \textit{kocht} “bought”, and \textit{after} instead of \textit{achter} “behind” were common in Holland and Utrecht.

Seventeenth-century examples of \textit{graft} instead of \textit{gracht} “canal”, \textit{koft} for \textit{kocht} “bought”, and \textit{saft} for \textit{zacht} “soft” are found in works by famous poets such as Bredero and Vondel, who lived in Amsterdam (Schönfeld/Van Loey 1954: 96). Commandeur (1988: 163–164), who did not find any \textit{ft} forms for nineteenth-century Amsterdam, did find some in eighteenth-century documents from Amsterdam, but only \textit{graft} “canal”, and mainly in toponyms such as \textit{Prinsegraft}. As \textit{ft} occurs frequently in his seventeenth-century sources (Commandeur 1988: 161), we would expect to find quite a few \textit{ft} forms, in particular in our seventeenth-century data.

5.3.1 \textbf{Seventeenth-century North Holland}

The persistence of \textit{ft} in seventeenth-century Holland is confirmed by our data. In the 39 letters of the seventeenth-century North-Holland subcorpus, we find instances such as \textit{verkoft} “sold” and \textit{gecoft} “bought” which occur in seven letters (18%), linked to Haarlem, Hoorn, Monnickendam and Edam. As reflexes of Gm. \textit{ft} are quite scarce, we decided to investigate all historical-phonologically possible positions in the entire seventeenth-century North-Holland subcorpus. In Table 24, we present the results of this study, with on the left-hand side the \textit{〈ft〉} spellings revealing localisable usage, and on the right hand side \textit{〈cht〉} and \textit{〈ght〉} spellings.

As can be seen in Table 24, only 12 out of 39 letters contain reflexes of Gm. \textit{ft}. As stated above, the grapheme \textit{〈ft〉} appears in seven of these letters. The total number of \textit{〈ft〉} spellings is 7, which amounts to about half of all historical-phonologically possible positions, viz. 15, in these letters. Furthermore, there appears to be a clear preference for one specific lexeme. In Table 25, we rearranged the data in accordance with the lexemes carrying the reflex of Gm. \textit{ft}. We use present-day standard spelling for the lexemes.

\footnote{T-deletion in final position is common throughout the Dutch language area; see Goeman (1999).}
Table 24. Frequencies of ⟨ft⟩ vs ⟨cht⟩ and ⟨ght⟩ in twelve letters from North Holland.

<table>
<thead>
<tr>
<th>Letter</th>
<th>⟨ft⟩</th>
<th>N</th>
<th>⟨cht⟩ or ⟨ght⟩</th>
<th>N</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>verkoft</td>
<td>1</td>
<td>nicht</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>koft</td>
<td>1</td>
<td>achter</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>vercoft</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>gecoft</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>gecoft</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>gecoft</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>gecoft</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>–</td>
<td></td>
<td>bekocht</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>–</td>
<td></td>
<td>verkoht</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>–</td>
<td></td>
<td>doorluchtiche</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>–</td>
<td></td>
<td>night</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>–</td>
<td></td>
<td>luchti, luchtig</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Total N</td>
<td>7</td>
<td>8</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total %</td>
<td>47</td>
<td>53</td>
<td>100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 25. Frequencies of ⟨ft⟩ vs ⟨cht⟩ and ⟨ght⟩ in twelve letters from North Holland.

<table>
<thead>
<tr>
<th>Lexeme:</th>
<th>kocht</th>
<th>lucht</th>
<th>nicht</th>
<th>achter</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>⟨ft⟩</td>
<td>7</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>7</td>
</tr>
<tr>
<td>⟨cht⟩ or ⟨ght⟩</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>15</td>
</tr>
</tbody>
</table>

The results suggest that ⟨ft⟩ is mainly restricted to the lexeme *kocht*, where seven out of nine instances (78%) are spelled ⟨ft⟩, while all other lexemes appear only with ⟨cht⟩. Note, however, that the other lexemes are much less frequent.

5.3.2 Seventeenth-century Amsterdam

A similar study as reported on in the previous section was carried out on the seventeenth-century Amsterdam subcorpus. We extracted all words with historical *ft*, and investigated the spelling used. Again, we found instances of ⟨ft⟩ representing localizable usage. The tokens appear in four out of 54 letters (7%), which is even less than the 18% in the seventeenth-century North-Holland subcorpus (Section 5.3.1).

In these four letters from the seventeenth-century Amsterdam subcorpus, there are 28 historical-phonological positions allowing for ⟨ft⟩. The grapheme ⟨ft⟩ is used in six instances (21%). As with *sk* in initial position, data are so scarce that conclusive evidence for social and gender distributions cannot be established, despite the fact that these analyses were initially planned in order to gain insight into such variable aspects.
What we do observe, however, is the minor role that localisable spellings come to play in the written code.

Table 26 sums up the results for the entire seventeenth-century Amsterdam subcorpus, divided per lexeme as this was shown to be instructive in the previous section.

Table 26. Frequencies of \textlangle ft\textrangle vs \textlangle cht\textrangle and \textlangle ght\textrangle in seventeenth-century Amsterdam.

<table>
<thead>
<tr>
<th>Lexemes:</th>
<th>kocht</th>
<th>nicht</th>
<th>achter</th>
<th>sticht</th>
<th>echter</th>
<th>-achtig</th>
<th>graft</th>
<th>Total N</th>
<th>Total %</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textlangle ft\textrangle</td>
<td>4</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>2</td>
<td>6</td>
<td>21</td>
</tr>
<tr>
<td>\textlangle cht\textrangle or \textlangle ght\textrangle</td>
<td>5</td>
<td>5</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>–</td>
<td>22</td>
<td>79</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>5</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>28</td>
<td>100</td>
</tr>
</tbody>
</table>

The localisable \textlangle ft\textrangle spelling appears to have been a fairly marginal phenomenon, not just when compared to the majority of \textlangle cht\textrangle and \textlangle ght\textrangle spellings, but also in terms of lexemes. As suggested by the North-Holland data (Section 5.3.1), the \textlangle ft\textrangle spelling mainly occurs in specific lexemes such as koft, where four out of nine tokens in the Amsterdam data (44%) actually have \textlangle ft\textrangle. Furthermore, another lexeme also seems to conserve \textlangle ft\textrangle, viz. graft, which corresponds to present-day standard Dutch gracht “canal”. Both instances are spelled \textlangle ft\textrangle. It should be noted that both instances appear in toponyms, viz. Prinse graft and nieuwe palm graft, recalling Commandeer’s (1988) conclusions on eighteenth-century Amsterdam, where \textlangle ft\textrangle was mainly found in similar toponyms.

5.3.3 Eighteenth-century North Holland

The situation in the eighteenth-century North-Holland subcorpus is very similar to what we have seen in the previous sections. In 9 out of the 24 letters making up the subcorpus, 28 historical-phonological positions are found where either \textlangle ft\textrangle or \textlangle gt\textrangle may occur (the spelling shift to \textlangle gt\textrangle is consistent throughout the eighteenth-century North-Holland subcorpus). Table 27 presents the date per lexeme.

As can be seen in Table 27, no more than five instances of \textlangle ft\textrangle are found, as opposed to 23 (or 82%) instances of \textlangle gt\textrangle. Note, furthermore, that both instances of graft refer to the town of Graft, which still carries that name today. We conclude that \textlangle ft\textrangle is very rare in eighteenth-century North Holland.

Table 27. Frequencies of \textlangle ft\textrangle vs \textlangle gt\textrangle in eighteenth-century North Holland.

<table>
<thead>
<tr>
<th>Lexemes:</th>
<th>kocht</th>
<th>nicht</th>
<th>achter</th>
<th>echt</th>
<th>hecht</th>
<th>graft</th>
<th>Total N</th>
<th>Total %</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textlangle ft\textrangle</td>
<td>2</td>
<td>–</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>2</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>\textlangle gt\textrangle</td>
<td>9</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>–</td>
<td>23</td>
<td>82</td>
</tr>
<tr>
<td>Total</td>
<td>11</td>
<td>9</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>28</td>
<td>100</td>
</tr>
</tbody>
</table>
5.3.4 Eighteenth-century Amsterdam

Localisable spellings such as ⟨ft⟩ being rare in seventeenth- and eighteenth-century North Holland (Sections 5.3.1 and 5.3.3) as well as in seventeenth-century Amsterdam (5.3.2), it is not surprising that ⟨ft⟩ is even much less frequent in the 166 letters of the eighteenth-century Amsterdam subcorpus. We found only one instance of ⟨ft⟩, viz. in the toponym Brouwers graft. We did not compare this single token with all historically-phonologically possible positions, that is, with all reflexes of Gm. ft. This would have been a somewhat excessive exercise, as the general picture is obvious: localisable spellings are simply very infrequent. Moreover, we cannot be entirely certain that [ft] was the usual Amsterdam pronunciation in the eighteenth century. Building on previous work (see the introduction to Section 5.3), [ft] was assumed for the seventeenth century, but for the eighteenth century no conclusive evidence exists. Our results so far do not exclude the possibility that [ft] had largely given way to [χt] in late-eighteenth-century Amsterdam. This is confirmed by our comparison of the two most frequent seventeenth-century ⟨ft⟩-lexemes, viz. koft “bought” and graft “canal,” with the occurrence of these lexemes in the eighteenth-century Amsterdam subcorpus. Both occur quite frequently in the subcorpus, albeit not with ⟨ft⟩, but with ⟨gt⟩. These 49 instances of kogt and 21 of gragt clearly represent supralocal, and probably by then also Amsterdam, usage. Again, the data suggest that towards the end of the eighteenth century, the change from ⟨ft⟩ to ⟨cht⟩ was nearly complete in Amsterdam.

5.4 A-like vowels in the seventeenth century

In this section, we focus on the interplay of phonology and orthography with regard to specific velar and palatal vowels, which are distinguished in many North-Holland dialects, though not in present-day standard Dutch, where they are all a-like. We will first discuss this feature in more detail, and then move on to a case study on the basis of the seventeenth-century North-Holland subcorpus.

5.4.1 The representation of a-like vowels

In present-day standard Dutch, we find water “water”, with a lengthened originally short vowel in open syllable, and schaap “sheep”, with an originally long vowel. Both originally different phonemes have merged into [a:] in present-day standard Dutch. Many North-Holland dialects, however, maintain a phonological opposition of [a:]-like pronunciations for reflexes of the Wgm. short vowel in open syllables, and palatal pronunciations such as [e:], [ɛ:] or [ɛI] for reflexes of (predominantly) originally long PGM ē. Thus, North Holland has been characterised as the water-skeep area (Kloeke 1933, 1934), where the second word has a palatal vowel, as in English.10

In the Middle Dutch period, both phonemes were usually not distinguished in writing, following the merger in many southern Dutch dialects (Schönfeld/van Loey

---

10. Skeep also has sk in initial position, see Section 5.2
1954: 91–92). Consequently, postmedieval supralocal writing systems based on southern writing practices did not distinguish the two phonemes which were commonly represented by the symbol ⟨a⟩, neither separately, for instance in open syllables, nor lengthened by ⟨e⟩ or ⟨a⟩. Consequently, we usually find what could be termed a-based graphemes, for instance:

- water “water”, and schaap or schaepe for “sheep” in the singular,
- and schaepen, schaapen or schaeopen for “sheep” in the plural.

Note that before r and especially before r+dental consonant, short as well as long a’s (mostly from Gm. short *e) are palatalised in many varieties of Dutch, resulting in variants such as aarde and eerde “earth”, and waarde and weerde “worth” (Schönfeld/Van Loey 1954: 65–66). These doublets have existed for a long time, with some of them continuing to exist in present-day Dutch.

Despite the fact that supralocal writing traditions normally did not distinguish between the historically different phonemes, localisable spelling differences have been attested. Kloeke (1934–35: 12–21) observes ⟨ee⟩ spellings for reflexes of PGM ē and other palatal vowels which sound [aː] in present-day standard Dutch, in seventeenth-century farces from North Holland, such as:

- gee instead of ga “go”
- versteen instead of verstaan or verstaen “understand”
- queet instead of quaat or quaet “angry”
- nee instead of na “after”
- leet instead of laet or laat “let”
- deer instead of daar or daer “there”.

On the other hand, Commandeur (1988: 161), discussing the seventeenth-century Amsterdam dialect and noting the phonological difference between lengthened a’s and palatal reflexes of PGM ē, claims that the difference was not systematically represented in the spelling owing to the lack of graphemes.

According to Kloeke, the ⟨ee⟩ spellings were used for stereotyping people from rural North Holland. Other characters in the farces were represented by the more common, supralocal ⟨ae⟩ spelling, suggesting dialectal and/or sociolectal differences in pronunciation. The data suggest that palatal pronunciations were common in the seventeenth century, as they still are in many North-Holland varieties. They, therefore, offer a good opportunity for investigating the interplay of phonology and orthography in letters from North Holland.

We extracted from the seventeenth-century North-Holland subcorpus all words with a-like vowels, that is, all words with lengthened [aː] in open syllable, and all words with reflexes of PGM ē, as well as a few other words with ⟨ee⟩ spellings where ⟨aa⟩ or ⟨ae⟩ would have been expected from a supralocal perspective. In the 39 letters in the subcorpus, ⟨ee⟩ spellings were found instead of a-based graphemes, thus revealing the
writers’ palatal pronunciation. The instances occur in letters from Hoorn, Enkhuizen and Monnickendam and include:

- *geet* instead of *gaet* or *gaat* “goes”, and *seet* instead of *saet* or *saat* “seed”.
- instances before *r* or *r+dental*, such as *deer* instead of *daer* or *daar* “there”, *neerden* “(the town of) Naarden”, *meert* instead of *maert* or *maart* “March”, and *steert* instead of *staert* or *staart* “tail”.

The examples provide solid evidence both that the palatal pronunciation existed in seventeenth-century North Holland and that letter writers used the alphabet for the orthographical representation of this localisable pronunciation. Consequently, writing practices are revealed of which the existence was unknown or uncertain until now. Apart from the commonly used variants \( \langle ae \rangle \) and \( \langle aa \rangle \), the lesser known variant \( \langle ee \rangle \) was in use, demarcating reflexes of PGM ē from lengthened *a* in open syllable.\(^{11}\)

This does not, however, mean that such \( \langle ee \rangle \) spellings were common, nor that the phonemic difference between reflexes of PGM ē and lengthened *a* in open syllable was systematically represented in the writing systems employed in the letters. As in the previous case studies, we examined the frequency of localisable \( \langle ee \rangle \) spellings in proportion to supralocal *a*-based graphemes in similar positions.

The relative frequency of \( \langle ee \rangle \) spellings was first established by studying the orthographical representation of all *a*-like vowels, that is, of lengthened *a*, originally long *a* (mainly derived from PGM ē), and *a* before *r+dental*. We furthermore distinguished between open and closed syllables. The overall results are in Table 28.

**Table 28.** Frequencies of \( \langle a \rangle \), \( \langle ae \rangle \), \( \langle aa \rangle \) and \( \langle ee \rangle \) in seventeenth-century North Holland.

<table>
<thead>
<tr>
<th></th>
<th>Lengthened <em>a</em></th>
<th>Originally long <em>a</em></th>
<th><em>a</em> before <em>r+dental</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Open</td>
<td>Closed</td>
<td>Open</td>
</tr>
<tr>
<td>( \langle a \rangle )</td>
<td>160</td>
<td>4</td>
<td>155</td>
</tr>
<tr>
<td>( \langle ae \rangle )</td>
<td>17</td>
<td>200</td>
<td>51</td>
</tr>
<tr>
<td>( \langle aa \rangle )</td>
<td>4</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>( \langle ee \rangle )</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>181</td>
<td>214</td>
<td>207</td>
</tr>
</tbody>
</table>

As becomes clear from Table 28, the number of \( \langle ee \rangle \) spellings is remarkably low. Only eight instances were found. Compared to the highest possible number of positions where \( \langle ee \rangle \) could have been used, viz. 858 (=207+607+44), this is astonishingly low. The *a*-based graphemes \( \langle a \rangle \), \( \langle ae \rangle \) and \( \langle aa \rangle \) dominate the results by far. The eight instances

---

\(^{11}\) Not all forms derive from PGM ē, for instance *deer* “there” has a vowel derived from PGM *a*. In many dialects, however, such vowels, especially before *r* or *r+dental*, fit in with reflexes of PGM ē, as *deer* amply demonstrates.
of ⟨ee⟩ were attested in six letters (out of 39, viz. 15%). In these six letters, there were no fewer than 174 instances of originally long vowels and vowels before r+dental, which means that the unambiguously localisable spelling ⟨ee⟩ was used in 8/174 or 5% of the highest possible number of instances.

The impression so far is that local usage gave way to supralocal usage. In the remainder of this section, we will focus on these a-like vowels results in order to reach a better understanding of the writing systems used, as these may have overruled localisable orthography.

5.4.2 The writing systems used

Table 28 shows that the originally long vowel and the vowel before r+dental were predominantly spelled ⟨ae⟩ in closed syllables, with 531/607 and 40/44 instances, respectively. This was also the common sign for the lengthened vowel in closed syllables (200/214). Conversely, the originally long vowel was written ⟨a⟩ in the majority of open syllables (155/207), as was the lengthened a in open syllables (160/181). In other words, instead of a phonological distribution of orthographical signs, there rather appears to have been a syllabic spelling system with predominantly ⟨a⟩ in open syllables and ⟨ae⟩ in closed syllables, regardless of the historical-phonological origin of the a-like vowels in question. We therefore reorganised the overall results in Table 29, combining all a-like vowels, and solely focusing on the syllable structure.

Table 29. Frequencies of ⟨a⟩, ⟨ae⟩, ⟨aa⟩ and ⟨ee⟩ in seventeenth-century North Holland.

<table>
<thead>
<tr>
<th></th>
<th>Open syllable</th>
<th>Closed syllable</th>
</tr>
</thead>
<tbody>
<tr>
<td>⟨a⟩</td>
<td>315</td>
<td>27</td>
</tr>
<tr>
<td>⟨ae⟩</td>
<td>68</td>
<td>771</td>
</tr>
<tr>
<td>⟨aa⟩</td>
<td>5</td>
<td>59</td>
</tr>
<tr>
<td>⟨ee⟩</td>
<td>–</td>
<td>8</td>
</tr>
<tr>
<td>Total</td>
<td>388</td>
<td>865</td>
</tr>
</tbody>
</table>

As becomes clear from Table 29, the common way of representing the a-like vowels in writing was by adopting a syllabic writing system, with ⟨a⟩ mainly in open syllables and ⟨ae⟩ as the dominant variant in closed syllables. This is corroborated by Figure 8, which shows the proportion of the four graphemes per syllable type.

12. Note that the vowels were first lengthened in open syllables, which were then closed, often as a result of schwa-apocope, e.g. Middle Dutch ane “on” with lengthened vowel, Modern Dutch aan after apocope, thus with lengthened vowel in closed syllable. Note also that in the seventeenth century, lengthening by doubling the original vowel, as in ⟨aa⟩, was a minor variant, which would, however, become the dominant spelling during the eighteenth century (see e.g. Vosters, Rutten & van der Wal 2010).
Figure 8 strongly suggests that the syllable type is decisive. In open syllables, ⟨a⟩ is used in 81% of all instances, whereas ⟨ae⟩ is used in 89% of all instances in closed syllables. This means that letter writers, who in their local dialects produced palatal vowels phonologically distinct from lengthened a’s, did not systematically represent this phonological difference in written Dutch. On the contrary, they used a syllabic spelling system corresponding to supraregional writing practices, which were in line with the phonological merger of the different a’s in what were at one time leading varieties such as dialects from Flanders, Brabant and South Holland (Schönfeld/van Loey 1954: 91). This is a clear case where “certain orthographic traditions originating in the southern Netherlands continued to influence the written language in Holland” (Howell 2006: 210), as a result of which dialectal distinctions were veiled.

![Graph showing proportions of ⟨a⟩, ⟨ae⟩, ⟨aa⟩ and ⟨ee⟩ in seventeenth-century North Holland.](image)

As stated above, Commandeur (1988: 161) referred to the phonological difference between lengthened a’s and palatal reflexes of PGM ē in seventeenth-century Amsterdam, while claiming that the difference was not systematically represented in the spelling owing to a lack of graphemes. Lack of graphemes is obviously not the case: examples such as geet “goes” and seet “seed” show that ⟨ee⟩ could very well serve for representing the palatal vowel. Furthermore, we could think of writing systems with for instance ⟨aa⟩ for lengthened a in closed syllable, and ⟨ae⟩ for the palatal vowel. Commandeur himself (1988: 161–162) points out that such a system was in fact in use in some contemporary literary texts. The fact that the phonological difference was not represented in the spelling cannot be attributed to a lack of graphemes; instead it is because writing systems were rather supralocal than local.
Still, the question arises whether the syllabic writing system attested for the overall results also holds for individual letters. Focusing on individual letters, the picture presented so far does not change. By way of illustration, the distribution in one letter is given in Table 30.

**Table 30.** Frequencies of ⟨a⟩, ⟨ae⟩ and ⟨aa⟩ in a letter from seventeenth-century North Holland.

<table>
<thead>
<tr>
<th>Lengthened a</th>
<th>Originally long a</th>
<th>a before r + dental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>Closed</td>
<td></td>
</tr>
<tr>
<td>⟨a⟩</td>
<td></td>
<td></td>
</tr>
<tr>
<td>eersame,</td>
<td>na (2), swager</td>
<td>aarde</td>
</tr>
<tr>
<td>gevaren</td>
<td>(2), genade</td>
<td></td>
</tr>
<tr>
<td>⟨ae⟩</td>
<td>aen (6x), plaets</td>
<td>verstaen (2),</td>
</tr>
<tr>
<td></td>
<td></td>
<td>saeijtijt, quaedt,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gaen, gestaen,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>naeste, daer, waerom,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>maer (2), jaer, waer</td>
</tr>
<tr>
<td>⟨aa⟩</td>
<td>caap</td>
<td>saat</td>
</tr>
</tbody>
</table>

In the letter presented in Table 30, the familiar syllabic pattern dominates, though there are occasional divergences. In closed syllables, the spelling ⟨aa⟩ is used twice, once for a lengthened a as in *caap* “cape”, once for a reflex of PGM ē as in *saat* “seed”. In all other cases, ⟨ae⟩ is used, regardless of the etymology of the word, as ⟨a⟩ is used in open syllables. The phonological difference is not represented in the orthography, and the divergences from the common pattern are probably not related to dialectal pronunciation, as they concern each phoneme only once. Note also that the originally long vowel of *saat* “seed” also appears in *saeijtijt* “sowing time”, but with ⟨ae⟩.

Another letter contains the distribution as in Table 31. The graphemes ⟨a⟩ and ⟨ae⟩ are regularly distributed over open and closed syllables, respectively, regardless of phonological differences. Exceptions are *quadt* “angry” and *neerden* “(the town of) Naarden”. As toponyms are often conservative, more striking than ⟨ee⟩ in *neerden*, are the common ⟨ae⟩ spellings for the probably equally palatal vowels in *daer* “there”, *maer* “but”, *laet* “let” and so on, not distinct at all from the lengthened *a*s in *aen* “on” and *maet* “mate” and so on.

The general picture of the overall results as well as individual patterns indicated in Tables 30 and 31 point to a supralocal spelling system where the syllable structure is decisive, rather than to locally oriented writing systems.

Nonetheless, we may perhaps not assume palatal pronunciation for all of the so-called originally long vowels. Moreover, as stated before (footnote 11), not all forms derive from PGM ē. Also, Weijnen (1966: 210–212) notes that present-day dialects exhibit a high degree of variation, depending on the following consonant. Apart from this, there are also areas in North Holland where the *a*s have in fact merged (Kloeke
Table 31. Frequencies of ⟨a⟩, ⟨ae⟩ and ⟨aa⟩ in a letter from seventeenth-century North Holland.

<table>
<thead>
<tr>
<th>Lenghtened a</th>
<th>Originally long a</th>
<th>a before r + dental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open</td>
<td>Closed</td>
<td>Open</td>
</tr>
<tr>
<td>⟨a⟩</td>
<td>geuaren, hale, beware, water</td>
<td>weldade, genade (2), verrader (2), verlate (2), slage, verslagen, quade, sware, genadich, salich (2)</td>
</tr>
<tr>
<td>⟨ae⟩</td>
<td>aen (10), bewaert, haeldt, maet</td>
<td>laet 4x, gedaen, gaen (2), verstaen, verlaet, staen, lichaem, daer (2), daerom, waer [V] (2), maer (6), haer, waer [ADV]</td>
</tr>
<tr>
<td>⟨ee⟩</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1933: 253). In other words, the phonological context and lexical diffusion may distort the results presented in the foregoing. Therefore, we will approach the same feature from one more angle, viz. the unambiguously localisable spellings which we did find, and build up the picture from there. When focusing on the unambiguous examples geet “goes” (found in Monnickendam) and seet “seed” (found in Hoorn), we could hypothesise that palatal vowels are mainly maintained before dentals. In order to test this, we investigated the fairly frequent forms of “goes” in all 39 letters of the seventeenth-century North-Holland subcorpus. The results are in Table 32.

Table 32. Spelling of the vowel in present-day standard (PDS) Dutch gaat “goes” in seventeenth-century North Holland.

<table>
<thead>
<tr>
<th>PDS gaat “goes”</th>
</tr>
</thead>
<tbody>
<tr>
<td>⟨ae⟩</td>
</tr>
<tr>
<td>⟨a⟩</td>
</tr>
<tr>
<td>⟨aa⟩</td>
</tr>
<tr>
<td>⟨ee⟩</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Again, ⟨ae⟩ is prevalent, and therefore the difference with lengthened a’s in closed vowels is obscured. Put differently, there is no evidence that a following -t promotes a
palatal realisation of the vowel. Focusing on the other forms, we note that gat with single ⟨a⟩ appears in a letter from Enkhuizen which almost consistently has ⟨ae⟩, except in gadt “goes”, but it also contains gaedt “goes” with ⟨ae⟩. The gat spelling also occurs in a letter from Edam, which additionally has lat “let” and quat “angry”, but also, and more often laet (twice), gaet (three times) and quaet. Finally, both tokens of gaat with ⟨aa⟩ are found in a letter sent from Hoorn, where ⟨aa⟩ is the normal grapheme for all a-like phonemes, thus for reflexes of PGM ē (e.g. hennipsaat “hemp seed”) as well as for lengthened a (e.g. maat “mate”). Summing up, it is highly improbable that any of the spelling differences are motivated by the urge to represent phonological differences in the written code, and even if the spelling differences result from such an urge, the phonological differences are still not systematically represented in writing.

We may conclude that a few marked attestations reveal the palatal pronunciation of a-like vowels, for instance geet “goes”, seet “seed”, deer “there”, and neerden “Naarden”. These forms illustrate that seventeenth-century North Holland was, as Kloeke said, a water-skeep area. This phonological difference is, however, not systematically represented in the written language, which is instead characterised by syllabic writing systems. Since localisable spellings of a-like vowels were so infrequent in the seventeenth-century North-Holland subcorpus, and since the studies of sk and ft in the Amsterdam subcorpora (Sections 5.2 and 5.3) did not lead to many new results, we refrained from investigating the a-like vowels in the eighteenth-century North-Holland subcorpus and in the Amsterdam subcorpora. We will therefore turn to discussing another variable feature, the morphological feature of prefix ge-, which, like the palatal vowel, has been used in seventeenth-century and eighteenth-century literature to stereotype rural or lower-class people.

5.5 Prefix ge- in the seventeenth century

The case study in the present section concerns the realisation of the ge-prefix in past participles. We will first discuss the feature is more detail, and then turn to the corpus results.

Past participles carry the prefix ge- in present-day standard Dutch, while in many present-day North-Holland varieties the reduced prefixes e- and ø- are employed. There is a distribution of both options, in that the prefix is absent (‘zero prefix’) in the most northern parts, while the more southern dialects of North Holland have e- (Hol 1941: 250). In many other varieties, however, present-day standard Dutch included, the past participle has the prefix ge-, as in gekomen “come”, geworden “become”. The corresponding North-Holland forms would be komen or ekomen, and worden or eworden.

Middle Dutch texts display zero prefixes mainly in participles of verbs with a perfective meaning, such as worden “become”, vonden “found”, komen “come”, bracht “brought” (Schönfeld/van Loey 1954: 156). Note that these are strictly speaking not reduced prefixes: these perfective verbs simply did not have a prefix at that time. Most other participles carried ge- in Middle Dutch texts, the use of which was extended to
perfective verbs with the effect that postmedieval written Dutch is generally characterised by ge-. In seventeenth and eighteenth-century literature, however, rural or lower-class people were often stereotyped by the (abundant) use of reduced prefixes (see Rutten 2008 for examples and discussion).

As reduced prefixes are fairly common in North Holland up to the present day, we decided to extract all past participles from the seventeenth-century North-Holland subcorpus in order to gain insight into the interplay of localisable morphology and supralocal writing traditions. Past participles with reduced prefixes were found in letters from Enkhuizen, for instance haeldt instead of gehaeldt “got”, and weest instead of geweest “been”. Note that in the letter with haeldt, another remarkable form occurs: sondt “healthy”, corresponding to present-day gezond. In this case, ge- is not the past participle prefix, but still it is analogically deleted, thereby confirming the weak status of the past participle prefix in the spoken language.

As with the a-like vowels (Section 5.4), however, these attestations do not at all imply that this localisable feature commonly occurred in the written language of North Holland. Of the 39 letters in the seventeenth-century North-Holland subcorpus, only three contain an instance of a past participle with a reduced prefix. In the other 36 letters (92%), only the supraregional variant with ge- appears, thus masking local morphology.

Next, we focused on the prominence of the reduced forms in the letters where they occur, vis-à-vis supralocal realisations on the past participle with ge-. As with the other features discussed in the present chapter, the degree of orality even decreased when we analysed in greater detail the three letters in which reduced prefixes were found. In all three letters, one single token of a reduced past participle prefix was attested, and in all three letters the reduced form constituted a small minority. The results are in Table 33.

Table 33. Frequency of reduced prefixes in three letters from seventeenth-century North Holland.

<table>
<thead>
<tr>
<th>Letter</th>
<th>o or ⟨e⟩</th>
<th>N</th>
<th>⟨ge⟩/⟨ghe⟩</th>
<th>N</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>haeldt</td>
<td>1</td>
<td>geuaren, geschrieuien, gedaen, gegroedt</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>weest</td>
<td>1</td>
<td>geuare(n) (3), Geschreuen, gedaen, gekomen (2), gehadt (2), geloopen, geweest (2), gegroedt</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>onteboden</td>
<td>1</td>
<td>gesprooken, ghevaeren, ghesgreven, ghescheven, ghekregen, ghelbeken [gebleken], ghesedt, ghedocht, ghetrocken, ghomen, ghehounden, gheeven (2), ofghekapt, opghehanghen, uiijtgheroeijt, ghemackt</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>Total N</td>
<td>3</td>
<td></td>
<td></td>
<td>34</td>
<td>37</td>
</tr>
<tr>
<td>Total %</td>
<td>8</td>
<td></td>
<td></td>
<td>92</td>
<td>100</td>
</tr>
</tbody>
</table>
As can be seen in Table 33, reduced prefixes are rarities in the written language from North Holland. They occur in only three letters in the subcorpus, in which letters only three tokens were found, representing 8% of the total number of past participles in these letters. Dialect morphology thus plays such a minor role here that we decided not to further examine reduced prefixes in the eighteenth-century North-Holland subcorpus nor in the Amsterdam subcorpora.

5.6 Long e’s in the seventeenth and eighteenth centuries

In Sections 4.5 and 4.6, we discussed the orthographical representation of etymologically distinct long e’s in the seventeenth-century Zeeland subcorpus and in the eighteenth-century Zeeland subcorpus. We referred to these two long e’s as the softlong ê and sharplong ê, respectively. From the sixteenth century onward, there was a supraregional writing tradition in which the softlong ê in open syllable was written ⟨ē⟩, while the sharplong ê in open syllables was spelled ⟨ee⟩. In closed syllables, both were commonly spelled ⟨ee⟩.

- **Softlong** ê developed through lengthening of the short vowels [ε] and [I]. Examples are the vowels in the first syllables of *leven* “live (verb)”, *rekenen* “count”, and *hemel* “heaven”. In closed syllable, an example would be *leef* “live (first person singular)”.
- **Sharplong** ê out of Wgm. diphthong *ai* is found in closed syllables in *steen* “stone” and *een* “one”, and in the first and open syllable in *heten* “be called”.

In Section 4.5, we saw that many letters in the seventeenth-century Zeeland subcorpus were characterised by a phonological distribution of ⟨ē⟩ and ⟨ee⟩ in line with the supraregional writing tradition. At the same time, we witnessed graphemisation of the writing system: the importance of phonology for orthographical practice was reduced, while choices directly linked to the written code became more important. This implies morphologisation and/or syllabification of the spelling, where morphological and/or syllabic aspects decide on the grapheme used, for instance uniformity of inflected forms and root forms, or uniformity of open and closed syllables.

Interestingly, a diachronic shift in preferred writing systems took place. In the eighteenth-century Zeeland subcorpus, only three letters were found to have a phonology-based distribution of ⟨ē⟩ and ⟨ee⟩, whereas about half of all the letters contained a syllabic system. The writing systems in use in the Zeeland subcorpora were summed up in Table 16, which we repeat here as Table 34.

The main system in the eighteenth-century letters from Zeeland was the syllabic system with ⟨ee⟩ in open syllables for softlong ê as well as sharplong ê. As a consequence, the phonological difference present in the spoken language is veiled. We hypothesised that this was an effect of convergence to the writing practice in North Holland, and in particular to Amsterdam, where the two phonemes had merged into one long e, probably already at the end of the sixteenth century.
Table 34. (= Table 16, Section 4.6). Writing systems in 72 letters from seventeenth-century Zeeland and in 25 letters from eighteenth-century Zeeland subcorpus.

<table>
<thead>
<tr>
<th></th>
<th>17th c.</th>
<th>18th c.</th>
<th>Total syllabic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonology-based</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N letters (Proportion)</td>
<td>31 (43%)</td>
<td>3 (4%)</td>
<td>7 (10%)</td>
</tr>
<tr>
<td>N letters (Proportion)</td>
<td>3 (12%)</td>
<td>2 (8%)</td>
<td>10 (40%)</td>
</tr>
</tbody>
</table>

In order to test this hypothesis, we randomly selected 50 letters linked to Amsterdam: 30 from the seventeenth-century Amsterdam subcorpus, and 20 from the eighteenth-century Amsterdam subcorpus. We limited ourselves to these 50 letters (instead of examining all letters in the Amsterdam subcorpora), since our main aim was to develop a reference point for the Zeeland results. We extracted all words with a long e in open syllable, distinguishing between the phonemes \( \tilde{e} \) and \( \acute{e} \), between the graphemes \( \langle e \rangle \) and \( \langle ee \rangle \), and between the seventeenth and the eighteenth century. The overall results of this study of letters linked to Amsterdam are given in Table 35.

Table 35. Frequency of \( \langle e \rangle \) and \( \langle ee \rangle \) spellings in 30 letters from seventeenth-century Amsterdam and in 20 letters from eighteenth-century Amsterdam.

<table>
<thead>
<tr>
<th></th>
<th>17th c.</th>
<th>18th c.</th>
<th>Total syllabic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Softlong ( \tilde{e} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>330</td>
<td>68</td>
<td>85</td>
</tr>
<tr>
<td>%</td>
<td>(68%)</td>
<td></td>
<td>(24%)</td>
</tr>
<tr>
<td>Sharp long ( \acute{e} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>157</td>
<td>32</td>
<td>273</td>
</tr>
<tr>
<td>%</td>
<td>(32%)</td>
<td></td>
<td>(76%)</td>
</tr>
<tr>
<td>Total</td>
<td>487</td>
<td>100</td>
<td>358</td>
</tr>
</tbody>
</table>

As in the letters to and from Zeeland (Section 4.6), the sharp long \( \acute{e} \) in open syllable is mostly spelled \( \langle ee \rangle \) in Amsterdam, in both the seventeenth and the eighteenth centuries. The pattern for soft long \( \tilde{e} \) also perfectly resembles the Zeeland results. See Figure 9, where we compare the choice of grapheme per century, based on Table 35. Figure 9 clearly shows the stable pattern for the sharp long \( \acute{e} \), recalling the equally stable pattern in Figure 2 in Section 4.6, concerning the representation of this phoneme in the Zeeland subcorpora.
Figure 9. Sharplong ē and softlong ē in seventeenth- and eighteenth-century Amsterdam.

The pattern for softlong ē also perfectly resembles the Zeeland results (see Figure 2, Section 4.6). Whereas it is written ⟨e⟩ in about two-thirds of all instances in the seventeenth century, it is conversely spelt ⟨ee⟩ in about three quarters of all instances in the eighteenth century. In Amsterdam, as in Zeeland, the previous phonologically oriented distribution gives way to a syllabification of the writing system, with ⟨ee⟩ for all long e’s in open syllable. Thus the development is similar in both regions.

Still, it is somewhat surprising that the Amsterdam data from the seventeenth century suggest at least in part a phonological distribution. The letters are from the 1660s and 1670s, whereas the merger of the phonemes occurred much earlier in the century, or even at the end of the sixteenth century. On the other hand, writers were able to use the supraregional phonology-based writing tradition despite the fact that they did not produce the phonological difference in their spoken language. To clarify these matters, we studied the writing system per letter, as we did for the Zeeland subcorpora (see Table 34), and saw major differences appear. From the 30 letters in the seventeenth-century Amsterdam subcorpus, 25 contained enough data to determine the writing system. From the 20 letters in the eighteenth-century Amsterdam subcorpus, 19 contained enough data for a decision on the writing system. The results are in Table 36. Remember that we also consider variation as a system (see Section 4.5).

A number of striking results are indicated in Table 36. Firstly, note that there are hardly any letters with a phonology-based system, despite the fact that we allowed for two exceptions (cf. Sections 4.5, 4.6). Secondly, note that the variable system occurs very often, in the seventeenth century in particular, and in any case more often than in

---

13. See Section 4.5 on how we determined writing systems here.
Table 36. Writing systems in 25 letters from seventeenth-century Amsterdam (30 letters in total) and in 19 letters from eighteenth-century Amsterdam (20 letters in total).

<table>
<thead>
<tr>
<th></th>
<th>Phonology-based</th>
<th>Syllabic</th>
<th>Morphological</th>
<th>Total syllabic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(e)</td>
<td>(ee)</td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td><strong>17th c.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N letters</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>(Proportion)</td>
<td>(8%)</td>
<td>(12%)</td>
<td>(12%)</td>
<td>(24%)</td>
</tr>
<tr>
<td><strong>18th c.</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N letters</td>
<td>-</td>
<td>1</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>(Proportion)</td>
<td>(5%)</td>
<td>(47%)</td>
<td>(53%)</td>
<td>(47%)</td>
</tr>
</tbody>
</table>

Zeeland. Furthermore, note the absence of morphological systems and the strong increase in syllabic systems.

What do these results mean, and how do they relate to the Zeeland results in Table 34? What we see in seventeenth-century Amsterdam is the effect of the phonemic merger. Despite the fact that the overall distribution presented in Table 35 suggested strong influence of phonology and/or etymology, hardly any individual letter employs the phonology-based system, while more than two-thirds show the variable system. At the same time, letter writers appear to have been looking for a new order, and so syllabic systems became more wide-spread, and even predominated, in the eighteenth century. It is unclear why there are no morphological systems attested, though it should be noted that syllabic systems are simpler than morphological systems which depend on analogy, thus on absent analogical forms.

The absence of morphological systems is important for the Zeeland results. The trend towards syllabification was clear from the results from this corpus. However, in Zeeland in the seventeenth century, there were also morphological systems in use that disappeared at a later stage. We conclude that writers from Zeeland converged to Amsterdam, where the merger of the phonemes first caused a high degree of variation, which was then partly levelled out by a syllabification of the system. Zeeland followed Amsterdam in the move towards syllabification, as well as in the reduction of morphological systems. Moreover, Zeeland converged to Amsterdam in the reduction of phonology-based systems. This is all the more striking, since there was in fact a sixteenth-century supraregional tradition based upon the historical-phonological difference. This means that the influence of Amsterdam must have been extremely strong. The convergence of Zeeland to Amsterdam is visualised in Figures 10 and 11, where we present the proportion of letters with a particular writing system per period, comparing seventeenth-century Zeeland with seventeenth-century Amsterdam in Figure 10, and eighteenth-century Zeeland with eighteenth-century Amsterdam in Figure 11.

Figure 10 demonstrates the strong differences between Zeeland and Amsterdam practices in the seventeenth century. Whereas the phonology-based system stands out
Figure 10. Writing systems for long e’s in seventeenth-century Amsterdam and Zeeland.

Figure 11. Writing systems for long e’s in eighteenth-century Amsterdam and Zeeland.
in Zeeland, variation prevails in the Amsterdam results, reflecting the merger of both long e’s in the spoken language.

Contrary to Figure 10, the patterns for Amsterdam and Zeeland are very similar in Figure 11, with the variable and the syllabic system ⟨ee⟩ as the dominant systems. The changing writing systems in Zeeland, and its eighteenth-century patterning in accordance with the Amsterdam results, provide solid evidence for the classic view of a supraregional variety that expands from Amsterdam into other regions. And this new supraregional practice was stronger, even in Zeeland, than the previous writing tradition, which did represent the phonological difference of the two long e’s.

5.7 Conclusions

Conclusions on the presence of unambiguously localisable phenomena in the letters to and from North Holland are similar to those drawn with respect to the letters to and from Zeeland. The dialect features discussed here usually occur in a minority of the letters, and letters containing the features mostly have only one or two tokens. Occasionally, letters contain many instances, which then make up the bulk of all attested instances.

Due to the low frequencies of localisable features it proved difficult to formulate conclusions in terms of social categories. In the case of sk, there were a few letters with relatively many tokens. It was striking that these letters were written by women. In Chapters 4 and 5, and in fact throughout the present book, we will argue that writing experience is an important – if not the most important – factor for the explanation of distributional patterns in the language of letters. Writing experience is a cover term for literacy, schooling opportunities, profession or occupation, referring to the extent to which people used the written language in their everyday lives. Writing experience was socially distributed as well as gendered, so with all possible reservations, we may say that it seems significant that sk mainly featured in letters by women.

Comparing the frequency of sk and ft in Amsterdam in the seventeenth and in the eighteenth centuries, a diachronic difference was established. The features were even less prominent in the eighteenth century. Again, we should point out that this change ties in with similar changes in writing experience, the increase of which characterises the eighteenth century (see Chapter 4).

Palatal vowels with an a-like counterpart in supraregional varieties were shown to be commonly spelled with supraregional a-based graphemes, such as ⟨a⟩, ⟨ae⟩ and ⟨aa⟩. Spelling choices were determined by syllable structures, rather than by localisable phonological differences. A similar syllabification of the writing system was observed in the use of the long e’s. We called syllabification and morphologisation of the writing system graphemisation, which implies that the written language acquires more supraregional elements, and to a lesser degree reflects localisable spoken varieties.

It seems strange that in the case of the a-like vowels syllabification of the writing system led predominantly to the monographeme ⟨a⟩ in open syllable, whereas the
digrapheme ⟨ee⟩ was preferred in the case of the long e’s in open syllable. Note, however, that ⟨ee⟩ in open syllable contrasts with ⟨e⟩ in open syllable in cases where schwa-pronunciation is needed. For instance, rendering softlong beleven “experience” [bəlevə(n)] as beleven, on the one hand obscures the etymological origin of the vowel in the second syllable, which is now spelled ⟨ee⟩, similar to sharp long ethereum, but on the other hand marks the difference with the schwa in the first, prefixed open syllable. As such, syllabification of the system of long e’s implies the representation of yet another phonological difference at the same time.

Summing up, the degree of orality of the letters, that is, the extent to which localisable oral features feature in the written code, is fairly low. Traces of spoken language notwithstanding, convergence to supralocal practice dominates the writing systems employed in the North-Holland and Amsterdam subcorpora. In the case of the long e’s, Amsterdam sets the trend of syllabification, which is followed by the Zeeland letter writers.

6. Conclusions

In the present chapter, we studied six stable localisable features: h-dropping, etymologically distinct long e’s, sk in initial position, Germanic ft, the prefix of the past participle ge-, and the phonological difference between a-like vowels and palatal vowels. The first two features were studied in letters to/from Zeeland, the other four were studied in letters to/from North Holland, and the long e’s, sk and ft were also studied in letters to/from Amsterdam. Our main research question was to what extent contain the letters traces of these localisable features, and as such of the spoken language? What is the degree of orality of the letters?

All in all, the degree of orality is low. We do find unambiguous traces of spoken language, which renders the letters valuable for historical-phonological and dialectological research. Still, the letters are not written in local dialect, and we find instead a considerable influence of supraregional convergence. In comparison with supraregional, nonlocalisable spellings, the dialect spellings constitute a minority. Rather than local spelling, we find supraregally orthographical levelling. Without getting caught in the trap of the standard language ideology, it should be concluded that in most cases, and in most letters, writers converged to supraregional writing practices, and they did so to a far greater extent than they represented local features in their writing. Even the many examples related to h-dropping, still a very stable element of dialects of Zeeland, constitute only a small minority in comparison to the overwhelmingly large number of supraregional spellings, despite the fact that the alphabet provided solid means to represent local dialect in the written language. Referring to the diverging opinions discussed in Section 3, we thus demonstrated the profound influence of supraregional writing systems: when people switched from spoken to written language, they often also switched from local to supraregional writing practices. Paraphrasing Mihm (1998; “intendiertes
Hochdeutsch”) and Vandenbussche (2002; “intended standard language”), we conclude that, in general, letter writers did not aim to write their local dialect, but instead employed an *intended supraregional variety*, and were, to a very large extent, perfectly able to do so.\(^{14}\)

Furthermore, in many cases only a few writers were responsible for a large number of the localisable spellings. This was clear, for instance, for *h*-dropping in the seventeenth and eighteenth centuries, as well as for *sk* in both centuries. Apparently, persistent dialect interference is a marginal phenomenon, linked to the practice of only some individuals. This even more strongly reduces the influence of the spoken language on the written code in general. It also draws attention to writing experience as a determining factor in the explanation of distributional patterns. We will take up the theme of writing experience in Chapter 4, and will return to it many times in the course of the present book.

---

\(^{14}\) The term “intended standard language or intended supraregional variety” characterises the intention of the letter writers. To avoid any misunderstanding, we stress that “intended” does not qualify the outcome as a ‘defective’ standard variety.
1. Introduction

The letters in our corpus exhibit a striking and intriguing characteristic, in terms of their formulaic language. This warrants our full attention and will be the topic of Chapters 3, 4 and 5. In the present chapter, we will discuss the wide variety of formulae commonly found in the seventeenth- and eighteenth-century letters. Here our goal is primarily descriptive. In Chapter 4, we will focus on variation and change in the use of formulaic language, and in Chapter 5 we will relate formulaic language to various aspects of the practice of letter-writing. The overarching theme throughout the three chapters is defined by the key question of this monograph, viz. what happens when people who were less involved in the written culture than, for instance, members of the nobility and non-noble ruling classes, sat down, took up their pen, and started to write? Did they rely on formulaic language? If so, to what extent? Is the variation in the use of formulaic language sociolinguistically patterned? Did the use of formulaic language change from the seventeenth to the eighteenth century?

It will shortly become clear what we mean by formulaic language (Section 2). Our approach differs from many modern approaches to formulaic language in that we focus on epistolary formulae. Whereas any string of frequently co-occurring words may be formulaic in the sense of Wray (2002), or may be considered a phrase or collocation in the sense of lexicological and phraseological approaches (e.g. Cowie 1998; Granger & Meunier 2008), or a construction or a pattern in the sense of construction grammar or pattern grammar (e.g. Goldberg 1995; Hunston & Francis 2000), we focus only on formulaic strings found repeatedly in letters, and, it may reasonably be assumed, even largely restricted to the language of letters.

In the present chapter, we will describe epistolary formulae found in seventeenth- and eighteenth-century letters, discuss their pragmatic functions, and identify their rhetorical meaning in the composition of a letter as a text. Ultimately, this implies describing the inventory of prefabricated, formulaic chunks of language that letter writers had at their disposal, and the position of these formulaic units in letters. Knowledge of these formulaic units provided letter writers with a context-independent framework into which the context-dependent pieces of information needed only to be inserted. In describing the formulae and discussing their functions, we primarily tie in with Elspaß (2005: 152–196), who discusses formulaic language in German nineteenth-century letters.
We will consistently indicate the relative frequencies of the formulae under discussion, and occasionally refer to diachronic differences, whilst in-depth studies of variation and change in the use of formulaic language will be presented in Chapter 4. We will make frequent references to counterparts of the epistolary formulae in other European languages in order to add a comparative perspective to the present chapter. However, we will not discuss all possible similarities to and differences from formulae used in neighbouring language areas, as this would require a fine-grained comparative analysis, preferably based upon primary sources. This would be beyond the scope of the present monograph. Similarly, we will not trace the long history of the formulae discussed, many of which date back as far as the late-medieval and Renaissance rhetorical art of letter writing, the *ars dictaminis*, and Latin and French models for business and legal writings, the *ars notaria* (see Nevalainen 2001, Poster & Mitchell 2007 and the references there).

In Section 2, we will briefly discuss a letter with a fair number of epistolary formulae so as to illustrate what we mean by formulaic language. In Section 3, we will introduce the subcorpora used for the present chapter. In Section 4, we will discuss the main functions of epistolary formulae, which will be related to text composition in Section 5. In Sections 6, 7 and 8, we will deal with the formulae most frequently found in the subcorpora. Section 9 will present the conclusions.

Figure 1. Address of the letter by Angenieten Cornelis.
2. A sample letter

In (1), we present a seventeenth-century autograph letter to illustrate the phenomenon of formulaic language, and of epistolary formulae in particular. Many variants of the formulae feature in this letter, and, although there are more formulae, which will be discussed extensively in the remainder of this chapter, this letter thus represents a reliable sample of the formulae found in our letter corpus.

The sample letter, which dates from 15 September 1664, is written by Angenieten Cornelis and addressed to her husband Roellandt IJosten Ostvorndijck. On the right-hand side, we indicate whether or not the different parts of the letter are (largely) formulaic.

(1) A seventeenth-century letter
1. te behandijgen aen roel
2. landt ijosten ostvorndijck
3. vaert op het schijp rotter
4. dam dat voert cappetein
5. lendert arijensen haechs
6. wandt uijt rotterdam met
7. vrijendt die godt bewaert
8. straetwaert ijn

9. looft godt e booven al dit geschreuen uit rotterdam
10. met alle vrient schaep schaep schaep den viftienden
11. september

12. eersaemen seer beminden man roellant iosten oost voren dick
13. en alderliesten man ick ul husvrou angenietge cornelis
14. laet ul weeten als dat ick en ul vaeder en min vaeder en
15. moeder noch klock en gesont sien en ick hoep min alderliesten
16. man roellant iosten oost vooren dick oock mede soo is waer het
17. anders het souden min niet lief om hooren weesen en ick
18. hoep dat het den goeden god godt langen sal laeten dueren
19. aen siel en aen lichaem ist ons salich ick hoep dat ghi met gesontheit
20. weeder thus komen mogen dat wij men kander weeder sien en
21. sprekken mogen min alder liesten man roellant ioosten oost
22. vooren dick ick wenssen ul veel hondert duisent mal
23. mael goeden nacht met een kus aen den brief
24. min pet oem een vrou is doet daer hebben ick mede te
25. begrauen gewest daer wort geseit dat wij buiet gelt
26. ont fangen sellen eersdaechs maer ick weet niet hoeveel
27. sij praeten dat ghi licht noch wel drie of vier maenden uit
28. weesen mocht ick senden ul hier twee niewen heemden
29. met een paer kusen\(^1\) met een kam met een mes met
30. drie nosdocken met een dassie ie met een sloepie
31. niet meer op dit pat danden heer bevolen min alder lieten
32. man ghi moet het min ten besten houden ick heb daer
33. geschreuen van dat goedt tebet stellen maer passchintge konden
34. niet nou moet ick het thuis houden hondert duisent
35. mael goeden nacht bij min
36. Angenieten Cornelis

**Line-by-line translation**

1. to deliver to Roel-
2. landt IJosten Ostvorndijck,
3. sails on the ship Rotterdam-
4. dam, that commands Captain
5. Lendert Arijensen Haechs-
6. wandt from Rotterdam, with
7. friend whom God protects
8. to the Mediterranean Sea

9. Praise God above all, this written from Rotterdam
10. with all friendship, the fifteenth
11. September

12. worthy very beloved husband Roellant Iosten Oostvorendick
13. and dearest husband, I, your wife Angenietge Cornelis
14. let you know that I and your father and my father and
15. mother are still in good health, and I hope my dearest
16. husband Roellant Iosten Oostvoorendick is also like that, if it were
17. different, I would regret to hear that, and I
18. hope that the good God will let it last a long time
19. to body and soul, we’d be blessed, I hope that you, with health,
20. may return home again, that we may see each other again and
21. speak to each other, my dearest husband Roellant Ilosten Oost-
22. voorendick, I wish you many hundred thousand
23. times good night with a kiss on the letter
24. my godfather’s wife is dead, I went to the
25. funeral, it is said that we a share of the loot
26. will receive one of these days, but I do not know how much
27. they say that you easily another three or four months
28. may be gone, I hereby send you two new shirts

---

\(^1\) Read *kousen* “stockings” with an omitted *u*. The alternative possible writing failure *kusens* “pillows” is less probable.
29. with a few stockings with a comb with a knife with
30. three handkerchiefs with a scarf with a pillow slip
31. nothing more on this occasion than commended to the Lord, my dearest
32. husband, you must not take it ill of me, I have
33. written to order the goods but Passchintge could
34. not make it, now I have to keep it at home, a hundred thousand
35. times good night, by me
36. Angenieten Cornelis

The address (lines 1–8) begins with the formula te behandijgen aen “to be delivered to”, after which the name of the addressee is filled in (Roellandt IJosten Ostvorndijck). This addressee’s identity is then detailed by indicating the ship he is working on, and again the name only needs to be filled in (the Rotterdam). Then, the ship is identified by its captain (Lendert Arijensen Haechswandt), as well as by its route. The address also contains the common formula met vrijendt die godt bewaert “[sent] with a friend who is under the protection of God”. These formulae will be discussed in more detail in Section 6.1. Here, we mainly want to point out the pattern in accordance with which the address is formulated. It consists of fixed phrases, interrupted only by context-dependent proper names. Similarly, the date (lines 9–11) consists of fixed phrases, and only the place and date need to be filled in.

Most letters follow the rhetoric of the sample letter in (1) in that they contain three main parts, viz. the introduction, the middle part and the ending. Of these, the introduction and the ending are usually highly formulaic and made up of context-independent fixed phrases. The middle part, on the contrary, commonly contains many more non-formulaic strings, expressing context-dependent information. At this point, we would like to draw attention to the fact that the extensive introduction of the letter (lines 12–23), conveying information on health and communicating the wish for renewed face-to-face contact, is no less formulaic than the preceding parts (address and date), and may also be considered as a pattern in which the letter writer only needs to fill in the appropriate names. Any single piece of information in the introduction, as personal as it may sound, can in fact be found in numerous other letters as well, as will be shown in Sections 6, 7 and 8.

The middle part (lines 24–30), where context-dependent information is conveyed such as deaths and financial and other housekeeping matters, does not contain phrases commonly found in other letters. In other words, the middle part does not contain any epistolary formulae. This does not imply that epistolary formulae never occur in the middle part (quite the contrary, see Section 6.2.) nor that the language of the middle part may not be idiomatic or may not consist of constructions, collocations or phrases (cf. Section 1).

The ending (lines 31–35) contains common epistolary formulae such as niet meer op dit pat [= pas “step”] than “lit. nothing more on this step than, nothing more for now but”, den heer bevolen “lit. the Lord commended, commended into the hands of the
Figure 2. Letter by Angenieten Cornelis, dated 15 September 1664.
Lord”, and *hondert duisent mael goeden nacht* “a hundred thousand times good night”. In between, the letter writer makes a non-formulaic, context-dependent remark about “the goods”. It often happens that these kinds of non-formulaic sequences are mixed with the otherwise formulaic ending. The reason must be that, while finishing the letter, important new information comes to the letter writer’s mind. As a result, the ending is not always as formulaic as the introduction, and it may even be difficult to decide where exactly the middle part stops and the ending begins (see also Section 6.1.4).

The language of the sample letter is highly formulaic. Thus, the letter sets a somewhat extreme example, which however clearly illustrates the importance of formulaic language for historical letter writing. Nonetheless, it should be borne in mind that there were also letters with fewer formulae, and that the distribution of formulae decreased over time and was subject to sociolinguistic effects (see Chapter 4), and that, finally, letters as formulaic as (1) may also have much longer non-formulaic middle parts.

3. The subcorpora

For the present study, we used two subcorpora, which together contain approximately 100,000 words. The basic external data are given in Table 1.

<table>
<thead>
<tr>
<th>Period</th>
<th>N letters</th>
<th>N words</th>
</tr>
</thead>
<tbody>
<tr>
<td>1660s-1670s</td>
<td>81</td>
<td>45,000</td>
</tr>
<tr>
<td>1780s</td>
<td>100</td>
<td>56,000</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>181</strong></td>
<td><strong>101,000</strong></td>
</tr>
</tbody>
</table>

In a pretest, it was found that after analysing ten letters from each period, new epistolary formulae appeared only sporadically. We may, therefore, safely assume that these 181 letters provide us with a representative sample of seventeenth- and eighteenth-century epistolary formulae. In Chapter 4, we will use larger corpora (cf. Chapter 1, Section 6) so as to be able to carry out sociolinguistic studies of formulaic language.

4. Functions of epistolary formulae

In this section, we present our approach to the functions of formulaic language in Dutch seventeenth- and eighteenth-century letters. First of all, we stress that the pragmatic situation in which the epistolary formulae acquire meaning consists of the texts (the letters) and the two participant roles of the writer and the addressee. Furthermore, we distinguish a third entity to whom reference is frequently made, viz. God.
Building on Elspaß (2005: 157–196), Wray (2002), Tannen (1987) and Rutten & van der Wal (2012), we distinguish three main functions of formulaic language, each foregrounding different aspects of this pragmatic situation: the text-constitutive, the intersubjective and the Christian-ritual function. These functions will be discussed in detail in Sections 6, 7 and 8, where extensive examples from the subcorpora are also presented. In the present section, we will introduce the three main functions, which will be related in Section 5 to the rhetoric of letter writing.

**Text-constitutive formulae**

Text-constitutive formulae foreground the text in itself, that is, they draw attention to the fact that the text is a letter. They express one of the main functions that formulae perform according to Wray (2002: 101), i.e. “marking of the discourse structure” or “discourse marking”. Two subtypes of text-constitutive formulae are distinguished, the first relating to the text type, the second to the text structure. Elspaß (2005: 157), discussing nineteenth-century German letters, speaks of “Textsortenkonstitution” and “Textkonstitution”, respectively. Laitinen & Nordlund (2012) describe similar functions in nineteenth-century Finnish correspondence.

The first subtype of text-constitutive formulae is made up of **text-type formulae**. Text-type formulae identify the text as a letter. When scrutinising the text for the first time, the text-type formulae may enable the addressee (or the researcher) to rapidly and easily determine that the document in question is a letter. The text-type formulae comprise address formulae, date formulae, salutation and opening formulae, closing formulae and signatures. Without actually having to read the body of the text, the text type is revealed by these surrounding text-type formulae. All the letters, apart from a few exceptions, contain these text-type formulae. Examples of text-type formulae in sample letter (1) are *te behandijgen aen* “to be delivered to” in the address, and *dit geschreuen uit* “this [has been] written in/from” in the date.

The second subtype of text-constitutive formulae is made up of **text-structural formulae** which mark the text structure by realising the transition from one part of the discourse to another. In principle, text-structural formulae can be used throughout the text, and, in fact, in any text. In practice, they are not only highly specific for letters, but also largely confined to specific text-structural functions. The first kind of text-structural formula marks the transition of the opening of the letter with its typical text-type formulae to the next part of the discourse, where the writer’s health is usually communicated. In (1), the recipient is addressed in lines 12–13. The transition from this opening phrase to the statement on the writer’s health in lines 14–15 is made by the text-structural formula *ick [ul husvrou angenietge cornelis] laet ul weeten als dat* “I [your housewife Angenietge Cornelis] let you know that”, where the part between the square brackets is context-dependent. Similar formulae are also frequently attested in Early and Late Modern English (Austin 1973a: 16–17). Whereas this *I let you know that* formula is usually found in the first part of the letters, near the beginning, another type
Chapter 3. Epistolary formulae

of text-structural formula commonly appears towards the end of the letters, marking the transition from the body of the text to the closing formulae. A fairly frequent example is found in (1), where this transition is made by the formula niet meer op dit pat [= pas “step”] than “lit. nothing more on this step than, nothing more for now but” in line 31. Again, English counterparts are attested by Austin (1973b: 129–130).

**Intersubjective formulae**

By focusing on the relationship between the writer and the addressee, intersubjective formulae foreground the interactional aspect of the pragmatic situation. They amount to what Wray (2002: 101) considers as one of the main functions of formulaic language: “the manipulation of the hearer (including the hearer’s perception of the speaker’s identity)” or in short “interaction”. In terms of content, intersubjective formulae cover three domains: health, greetings and contact.

Intersubjective **health formulae**, which are also well-known from the history of English (Davis 1965; Austin 1973a, 2004; Nevalainen 2001), consist of health statements and health wishes, and may add up to a string of 4, 5 or even 6 successive formulae. In (1), the writer communicates her own health and that of her parents as well as the recipient’s father’s health in lines 14–15. In lines 15–16, she hopes that the recipient is still healthy as well, adding in lines 16–17 that she would regret it if this were not the case. Finally, in lines 17–18, she hopes that God will make this state of good health last for a long time. The formulaic strings in lines 15–18 all follow the initial health statement in lines 14–15, and, generally speaking, these formulae always follow such an initial health statement, and hardly ever appear separately.

Intersubjective **greeting formulae** display a wide range of variation (Nevalainen & Raumolin-Brunberg 1995). A first subdivision is made between greeting formulae addressed to the addressee and greeting formulae addressed to a third party, where the addressee functions as intermediary. A second subdivision is made concerning the kind of wish expressed. Cases where the writer simply greets the addressee or bids him or her good night are the two main types. In (1), lines 22–23 convey a greeting formula, where the writer bids the addressee a hundred thousand times good night.

Through the use of intersubjective **contact formulae**, writers express the wish that epistolary contact with the addressee will be maintained or renewed, and/or the wish that contact in person will take place in the future. In (1), one of the most frequent contact formulae is in lines 20–21: dat wij men kander weeder zien en spreken mogen “that we may see and speak to each other again”.

**Christian-ritual formulae**

The third main function of epistolary formulae in seventeenth- and eighteenth-century Dutch letters foregrounds the relationship between the writer and the divine world, or between the writer, the addressee and the divine world. The Christian-ritual formulae
usually place the writer and/or the addressee under divine protection, thereby construing the writer’s identity as religious. Thus, the writer’s stylisation of the self (Coupland 2007) involves ethical reliability. In Wray’s (2002) perspective, Christian-ritual formulae might be termed a subset of interactional or intersubjective formulae, particularly since these include formulae that promote (the hearer’s perception of) the speaker’s identity, including “being perceived as a full member of whichever groups are deemed desirable” (Wray 2002: 96). The abundance of Christian-ritual formulae, however, and the fact that God or the divine world can be described as separate entities in the Early Modern pragmatic situation, leads us to define the Christian-ritual formulae as fulfilling a function of their own in epistolary discourse. The most frequent Christian-ritual formula is the commendation formula, with which the writer commends the addressee into the hands of God. In (1), this formula is found in line 31.

It is important to note that these three main functions do not constitute three separate categories of formulae. Rather, formulae may combine two or three functions, with one function being dominant. Moreover, all epistolary formulae are text-constitutive in that they only or mainly appear in letters and thus mark this specific text type. Wray (2002: 88) rightly comments “that any given expression can have more than one function at the same time”. Especially at the beginning and the end of letters, formulae may appear that combine the text-constitutive and the intersubjective function. At the beginning, after the (text-constitutive) address and the (text-constitutive) date, a transition to the intersubjective statements in the introduction needs to be established. A frequent formula used for this transition, which does not however feature in (1), is een vriendelijke groet en zij geschreven aan “a friendly greeting be written to” (see Section 6.1.3), after which the (context-dependent) name of the addressee should be filled in. This formula opens up the discourse; it places emphasis on the fact that the greeting given is written, and hence on the fact that the document in question is a letter. As such, the formula fulfils a text-constitutive function. At the same time, it represents a greeting to the addressee, and by filling in his name after the preposition aen “to”, the intersubjective aspect of written communication is confirmed. On the one hand, the formula closes the preceding text-constitutive parts (address, date), and on the other hand, it opens up the intersubjective perspective.

Similarly, the intersubjective and the Christian-ritual function may be combined. This mainly occurs when intersubjective contents are confirmed or reinforced through an invocation of God, or when they are put in religious or ritualised terms. An example is the afore-mentioned commendation formula, with which the writer commends the addressee into the hands of God (line 31 in (1)). This is a Christian-ritual formula, putting all earthly matters into a religious perspective. At the same time, it expresses the writer’s wish that the addressee be safe and sound, thus clearly adding an intersubjective element to the Christian-ritual perspective.

Finally, a formula such as looft goedt u booven al “praise God above all, before everything else” in line 9 in (1) is considered a text-constitutive formula because it only
appears in dates. At the same time, it makes reference to God and clearly adds the Christian-ritual function to its text-constitutive function.

5. Text composition

Having described the three main functions of epistolary formulae, we now turn to the place occupied by these formulae in the letter. In general, the body of a letter consists of three parts: an introduction, a middle part and an ending. The tripartite body of a letter is embedded in text-type formulae marking the beginning and the ending of the letter, respectively. Letters tend to begin with an outside address, an inside address, a salutation and an opening phrase. The date is placed either among these preliminaries or at the end of the letter, amidst rounding off parts such as the closing formula and the signature. These text-type formulae enable the addressee to determine the genre of the text he or she is holding as well as the most important external data: who wrote this letter, to whom, when and from where?

After the preliminaries, the introduction follows as the first full part of the body of the letter. Likewise, the closing formula and the signature are preceded by the third part of the letter, viz. the ending. The introduction and the ending are dominated by intersubjective formulae. In these passages, information on health and greetings are communicated, larded with Christian-ritual formulae. The most commonly used Christian-ritual formula, the commendation formula, is usually found in the ending of the letter. Text-structural formulae are used for linking the different parts of the discourse, and they also mark the boundaries between them. The highly formulaic introduction and ending embed the predominantly non-formulaic middle part of the letter, where context-dependent information is communicated. The middle part may contain text-structural formulae, however, such as *ik laat u weten als dat* “I let you know that”.

Table 2 presents the common structure of a letter. From top to bottom, a letter contains preliminaries such as the address, the date and the salutation. This formulaic part is followed by the equally formulaic introduction, containing mostly health formulae. The second part of the body of a letter is the, in principle, non-formulaic middle part with context-dependent information, which may, however, incorporate text-structural formulae. The length of the middle part varies from one or two short sentences (as in example (1)) to several paragraphs. Then comes another formulaic part, the ending, which usually contains intersubjective and Christian-ritual formulae. The letter closes with a closing formula and the signature, sometimes with the date. The formulae in Table 2 are English translations of epistolary formulae which will be discussed in the following sections.

---

2. In the seventeenth century, the date is found predominantly at the top of letters, close to the salutation. In the eighteenth century, the date is more often placed at the end of the letter, close to the signature; see Section 6.1.2.
Table 2. Common structure of a letter and the corresponding epistolary formulae.

<table>
<thead>
<tr>
<th>Common structure of a letter</th>
<th>Formulaicity</th>
<th>Type of formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Address</td>
<td>Largely Formulaic</td>
<td>- Text-type</td>
</tr>
<tr>
<td>- Date</td>
<td></td>
<td>- Text-type</td>
</tr>
<tr>
<td>- Salutation/opening</td>
<td></td>
<td>- Text-type</td>
</tr>
</tbody>
</table>

1. Introduction
- I hereby let you know that
- that your father, my father and mother and I are still in good health
- For the rest, I’m letting you know that

2. Middle part
- context-dependent information
- may contain e.g. For the rest, I’m letting you know that or as far as I’m concerned

3. Ending
- I wish you all the best
- I send you my kindest regards
- Be commended in mercy into (the hands of) God
- that we may see each other again with love
- Nothing more for now but

6. Text-constitutive formulae

As stated above, we distinguish between two types of text-constitutive formula, viz. text-type formulae (6.1) and text-structural formulae (6.2).

6.1 Text-type formulae

There is a wide range of text-type formulae identifying the text as a letter. We distinguish between the address (6.1.1), the date (6.1.2) and the opening and closing parts (6.1.3, 6.1.4). The opening and closing parts link the text-constitutive parts to the body of the letter, in particular to the formulaic intersubjective parts of the letters, and as such there is a clearly intersubjective element in the opening and closing formulae.
6.1.1 Address
The letters’ addresses are conventionalised and usually consist of five elements at the most:

1. a prepositional phrase having the meaning “to”
2. the addressee’s name and social role
3. the addressee’s location
4. the means by which the letter is sent
5. additional information.

The second and third elements are commonly found in any address. All five elements are highly formalised, which will be demonstrated in the following subsections.

Prepositional phrases having the meaning “to”

The addressee’s name is often preceded by a prepositional phrase having the meaning “to”. The shortest formula in use is simply the preposition aen/aan/an “to”. Occasionally, we find the French version a, or a different preposition altogether, viz. omme “lit. around, to”. In the seventeenth century, in particular, another formula was in use: dese brief sal men bestellen aen “this letter has to be delivered to” and its variants (2a–g). (2a–c) are full formulae, differing in syntactic structure (2a vs 2b, c) or in the main verb (2a, b vs 2c); (2d–g) are reduced variants. These reduced variants may maintain either the verbal part of the full formula (2d, e) or the nominal part (2f, g). All formulae end with the preposition aen/aan “to”.

(2) a. Dese brief sal men bestellen aen
   this letter shall one deliver to
   b. Deesen brijef te bestelle aen
   this letter to deliver to
   c. dese brief te behandige aan
   this letter to deliver to
   d. Te bestellen aen
   to deliver to
   e. te behandijgen aen
   to deliver to
   f. desen bryef aen
   this letter to
   g. dese aan
   this to

Apart from a single preposition, or a multi-word unit as in (2), a third option was the zero option of leaving this first element of the address blank, and starting the address
with the addressee’s name. We investigated the frequencies of these three options for each period; see Table 3. Note that we do not have an address for every letter in the subcorpora; for instance, of the 81 letters in the seventeenth-century subcorpus used for the present chapter, we have 71 addresses.3

Table 3. Prepositional phrases meaning “to” in addresses in the seventeenth and eighteenth centuries.

<table>
<thead>
<tr>
<th>Period</th>
<th>N letters</th>
<th>N addresses</th>
<th>N preposition</th>
<th>N multi-word</th>
<th>N ø</th>
</tr>
</thead>
<tbody>
<tr>
<td>1660s-1670s</td>
<td>81</td>
<td>71</td>
<td>37</td>
<td>22</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>(100%)</td>
<td>(52%)</td>
<td>(31%)</td>
<td>(17%)</td>
<td></td>
</tr>
<tr>
<td>1780s</td>
<td>100</td>
<td>85</td>
<td>31</td>
<td>7</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>(100%)</td>
<td>(37%)</td>
<td>(8%)</td>
<td>(55%)</td>
<td></td>
</tr>
</tbody>
</table>

There is a clear diachronic difference, in that the use of formulaic expressions, either a simple preposition or a multi-word string, decreases over time. At the same time, the zero option increases from 17% in the seventeenth century to 55% in the eighteenth century. In Chapter 4, we will show that there is a general decrease of formulaic language from the seventeenth to the eighteenth century.

Finally, as letters were often bundled, and intended to be passed on, formulae similar to those in (2) were in use to indicate the actual addressee (3a–c).

(3) a. \( \text{om voort te bestellen aen} \)
    to further to deliver to
    ‘to deliver further to, to pass on to’

b. \( \text{om voort te behandige aen} \)
    to further to deliver to

c. \( \text{te Bestelle aen} \)
    to deliver to

The addressee’s name and social role

The addressee’s name, the key information of the address, usually consists of both the first and the last name and can be modified in different ways. A specification of the addressee’s social role is also very common, that is, either his profession or his relationship to the sender. Nevala (2004b: 86–92), discussing early English correspondence,

---

3. Some addresses lacked in the letters, others, written on separate sheets or envelopes, were lost in the course of time. Note that our late-eighteenth-century letters were often sent in envelopes, while most seventeenth-century letters were folded and provided with an address written on the outside of the letter itself.
refers to these specifications as occupational terms and terms of kinship/friendship, respectively.

As stated above, we found 71 addresses in the seventeenth-century subcorpus used here. 79 proper names occur in these 71 addresses (recall that letters were often passed on to someone else). In these 79 proper names, no less than 46 of such functional specifications are found, nine of which appear before the proper name (4a), and 37 following it (4b).

(4) a. Aen schipper Waligh Jans keijser
to shipmaster Waligh Jans Keijser

b. aen den eersamen janman lourens verbeeck botte lier
to the honoured crewmember Lourens Verbeeck cupbearer

Another way of modifying the addresssee’s name is by adding an adjective. Focusing, again, on the 79 proper names in 71 addresses from the seventeenth century, we find 31 proper names modified by an adjective, for which 41 adjectives are used. The most commonly used adjective is *eersame* “honoured” (21 tokens), see (4b). Other examples include *vrome* “pious”, *discrete* “discrete”, *eerbare* “honourable”, *voorsienige* “prudent”. There is a striking parallel with the adjectives used in early English correspondence, where *right*, *trusty*, *honoured*, *good*, *reverend* and *beloved* are the most typical examples (Nevala 2004b: 91), as well as with the use of adjectives such as *wert* “worthy” and *teuer* “dear, beloved” in German nineteenth-century letters (Elspaß 2005: 160).

Thirdly, the addresssee’s name can be preceded by a general term of address, such as *mijn heer* “Mr, Sir,” (4c) or its French counterpart (4d). These are, obviously, similar to *Master, Mistress, Lord, Lady, Sir, Madam* and the like, commonly used in English letters of earlier times (Nevala 2004b: 91).

(4) c. Mijn heer Mijn heer jacobus Veth
Mr Mr Jacobus Veth

d. Monsr Monsr Louwrent Willemsen Verpoorte
Mr Mr Louwrent Willemsen Verpoorte

As can be seen in (4c,d), these general terms of address are often used twice, with the first instance graphically separated from the rest of the address. The different ways of modifying the addresssee’s proper name may naturally occur simultaneously, as can be seen in the eighteenth-century example in (4e).

(4) e. Wel Eerw en zeer geleerde Heere Dhra J. van der Linden
Well honoured and very learned sir Mr J. van der Linden
rusted Predikant
emeritus pastor
The addressee’s location

The addressee’s name is usually followed by his location. The transition to this new piece of information is not created by graphic means such as placing the location on the next line, but by prepositions. Figure 3 presents an example. It gives the address of a seventeenth-century letter, but instead of, for instance, placing the name and the location (a ship) on different lines, as modern Dutch conventions would require, the different elements of the address are attached to each other by prepositions. The transcription and the translation are given below Figure 3.

Figure 3. Address of Grjetje Jans’ letter to Seewert Leenders 30 May 1672.

Line-by-line transcription and translation

\[
\begin{align*}
te \ behandigh & \ an \ seewert \\
\text{to deliver} & \text{ to Seewert} \\
leenders & \text{ ondersturman} \\
\text{Leenders second mate} & \\
op \text{ het schep de Jonge} & \\
\text{on the ship De Jonge} & \\
prjns \ [... \] & \text{ met vrjnt} \\
\text{Prins [The young prince] [...] with friend} & \\
die & \text{ godt bevaert} \\
\text{whom God protects} & \\
\end{align*}
\]

Other examples are in (5a–d). The prepositions introducing the location are in boldface.
(5) a. Aen leintje frans huijs vrouwe van lieven de wever in de korte
to Leintje Frans wife of Lieven de Wever in the Kort
noort strate Tot Vlissinge
Noortstrate in Vlissingen

b. Aan d Huysvrou van Cornelis Remmers In de bekkerstraad a
to the wife of Cornelis Remmers in the Bekkerstraat in
Amsterdam
Amsterdam

c. Mijn Heer De Hr D: van Horsen coopman bij de Maarepoort tot
Mr Mr D. van Horsen merchant near the Maarepoort in
Leijden
Leiden

d. Aan Mathys Artenaas Wonende Int Zeemans Welvaren op de
to Mathys Artenaas living in the Zeemans Welvaren in the
dokke Te Vlissingen In Zeland
dock in Vlissingen in Zeeland

It should be noted that the location need not be the name of a street as in (5a, b), but
may also be a ship (Figure 3) or a landmark (5c). Note also that before place names tot
“in” is used in the seventeenth century (5a), whereas there is variation of tot and te in
the eighteenth century (5c, d). The French counterpart a was also in use (5b).

The transition to the location may also be established by a verb phrase. The most
frequently used verb phrase is the present participial phrase (5e, f), but past participial
phrases (5g) and present tense phrases (5h) are used as well. The most frequently ap-
pearing verb is wonen “live” (5e, f, h), but other verbs such as varen “sail” (5g) are also
used, dependent on the addressee’s actual location ashore or on board ship.

(5) e. dese brief te bestellen aan de Weduwe van johannes Weber
this letter to deliver to the widow of Johannes Weber
woonnen de op de laan tot schiedam in hollandt
living on the street to Schiedam in Holland

f. Aan D’ Eerzaame Huijsvrouw Van Capt’n: Jan de Groot Jansz
to the honoured wife of captain Jan de Groot Jansz
Woonende op de linde gragt, tenhuijze Van Cap’n. H:
living on the Lindegragt, at-the-house of Captain H.
Rave, a Amsterdam
Rave, in Amsterdam

g. Mijn heer mijn heer gerard vander meer om voort te bestellen
Mr Mr Gerard van der Meer to further to deliver
dirck pranger uijt gevaren met het wapen van alckmaer in
Dirck Pranger sailed away with Het Wapen van Alckmaer in
Letters as Loot

t. jaer 1672 Tot Batavia
the year 1672 to Batavia

h. mijn heer mijn heer martinus caesar gewesen opperhoofdt in japan
Mr Mr Martinus Caesar former chief in Japan

woont op de tijgers graft op batavia
lives at the Tijgersgraft in Batavia

Finally, we find a few examples where an adjective establishes the link between the addressee’s name and location, viz. woonagtig “living”, which is derived from the verb wonen “live” (5i).

(5) i. Aan Hiltje Jacobs Gucksen Woonagtig op De Roosze Gragt A
to Hiltje Jacobs Gucksen living at De Rooszegragt in
Harlingen in Vriesland
Harlingen in Friesland

The means by which the letter is sent

In the seventeenth as well as in the eighteenth century, the means by which the letter is sent is indicated in over half of the addresses: in the seventeenth century, in 41 out of 71 addresses (58%), and in the eighteenth century, in 50 out 85 addresses (59%). The most common means in the seventeenth century (32 tokens) are variants of the formula met vrient die godt bewaert “[sent] with a friend who is protected by God” (6a). In the eighteenth century, a different verb is usually employed (6b). The formulaic nature of the expression is exemplified by reduced forms (6c) and abbreviated forms (6d). French versions (6e) and mixed Dutch-French versions (6f) are also found.

(6) a. met vrient die godt bewaert
with friend whom God protects
b. met vrint die godt geleijt
with friend whom God leads
c. met vrint
with friend
d. Met V. d: G: geleijde
with friend whom God leads+SUBJ
e. per Amis
with friend
f. par vrindt
with friend

In the eighteenth century, in particular, the formula is often preceded by a specification of the person who is actually carrying the letter (6g–i). This person’s name can be
modified (6h) as explained above, and is commonly preceded by the preposition per. What is also interesting, is that the reduced formula D.G.G., an abbreviation of dien God geleid “whom God leads”, is the most common variant in the eighteenth century. Frequency of use has reduced the form of the formula over time.

(6) g. P: Capt: E Redouw a Amsterdam D.G.G. through captain E. Redouw to Amsterdam whom-God-leads

h. P Capteijn Swerus Vrolijk Commandeerende ’t OJ through captain Swerus Vrolijk commanding the East India Comp: Schip De Heldt Woltemade D: G: G: Company ship De Heldt Woltemade whom-God-leads

i. Per Capt. de heyde dGG o Zee. through captain De Heyde whom-God-leads over sea

Note, finally, that over zee “over sea” is sometimes added, which may also be abbreviated (6i).

Additional information

Additional information only sporadically follows the address as outlined so far. In the seventeenth-century subcorpus, this happens only five times, in the eighteenth-century subcorpus only three times. Additional information mostly concerns the date (7a) or epistolary idioms (7b, c).

(7) a. anno 1673 ii³ 30 in the year 1673 February 30

b. cito cito cito with dispatch, with dispatch, with dispatch

c. in salvo in safety

6.1.2 Date

Like the addresses, the letters’ dates are highly conventionalised. They usually consist of four elements at the most, of which the actual date is one. The second and third elements appear in most letters.

1. praise to God
2. the writer’s location
3. the date
4. additional information.

These elements will be discussed in the following subsections.

4. ii “two”, indicating the second month of the year.
**Praise to God**

The first element almost only appears in seventeenth-century letters, where 35 out of 82 letters of the subcorpus (43%) contain the first element. The common variant is *looft godt boven al* “praise God above all” (8a), which appears 21 times, amounting to 60% of all cases. Variants include reduced forms (8b) and Latinised forms (8c), which both appear several times, as well as occasional variants (8d,e).

(8) a. *looft godt boven al*
praise God above all

b. *looft godt*
praise God

c. *Laus deo*
praise to-God

d. *Loft godt altijt waer dat geij sijt*
praise God always where that you are

e. *looft godt bouen alle menschen*
praise God above all people

The eighteenth-century subcorpus used for the present chapter does not contain a single example of this formula. In fact, in the entire eighteenth-century corpus used for the present book, only one instance is found, viz. in a letter dated 11 November 1780 by the lower-middle class woman Catarinna Cahttou. She places *Loeft god Bo-even al* “Praise God above all” at the beginning of her letter, while the place and the date are at the end so that the praise to God is separated from these elements.

**The writer’s location**

In the seventeenth century, the writer’s location, as a part of the date, is most often placed at the beginning of the letter, as in the sample letter in (1). In the eighteenth century subcorpus, more locations are found at the end of the letters, which is line with a general trend towards dating at the end (Simons 2009: 176). As can be seen in Table 4, 43 out of 63 dates (68%) in the seventeenth-century selection used here are placed at

<table>
<thead>
<tr>
<th>Period</th>
<th>N letters</th>
<th>N locations</th>
<th>N beginning</th>
<th>N end</th>
</tr>
</thead>
<tbody>
<tr>
<td>1660s-1670s</td>
<td>81</td>
<td>63</td>
<td>43</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(100%)</td>
<td>(68%)</td>
<td>(32%)</td>
</tr>
<tr>
<td>1780s</td>
<td>100</td>
<td>91</td>
<td>40</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(100%)</td>
<td>(44%)</td>
<td>(56%)</td>
</tr>
</tbody>
</table>

Table 4. Place of the writer’s location in the seventeenth and eighteenth centuries.
the beginning, while the balance is slightly in favour of dates being placed at the end of letters in the eighteenth century, when 51 out of 91 dates (56%) occur at the end.

In the seventeenth-century subcorpus, place names are usually preceded by a preposition (55 out of 63 locations, or), usually in (9a), but in a few scattered instances other prepositions appear (9b–e).

(9) a. in rotterdam
    in Rotterdam
b. t amsterdam
    at Amsterdam
c. uit rotterdam
    from Rotterdam
d. binnen Sername
    in Surinam
e. ady egmont op zee den 10 novemb Ano 1672
    lit. anno Domini, in Egmond aan Zee the 10 November Anno 1672

The preposition ady in (9e) is a special case since it was originally an abbreviation of anno Domini. It has been reinterpreted as meaning “in”, as will be explained in the following subsection (see (10i)). A remarkable difference in the eighteenth-century subcorpus is the complete absence of prepositions before place names. In a few instances, the writer’s location is preceded by the phrase geschreven “written” (plus preposition) or its Latin counterpart actum; see (9f, g).

(9) f. geschreven uijt amsterdam
    written from Amsterdam
g. Actum in rotterdam
    made in Rotterdam

Of such cases, 11 occur in the seventeenth-century subcorpus and 3 in the eighteenth-century subcorpus, where obviously the place name is preceded by a preposition in these instances.

The date

As the writer’s location moved from predominantly at the beginning of the letters to mainly at the end of the letters, likewise the date itself is more often placed at the end of the letters in the eighteenth-century subcorpus (56%) than in the seventeenth-century subcorpus (27%); see Table 5.

The common pattern in the seventeenth and eighteenth centuries is den-day-month-year, where den “the” is the accusative singular of the masculine definite article (10a, b), and where especially the article or the year is sometimes left out (10c, d). Incidentally, the year is put first (10e). Occasional variants include de “the” (nom sing
Table 5. Place of the date in the seventeenth and eighteenth centuries.

<table>
<thead>
<tr>
<th>Period</th>
<th>N letters</th>
<th>N dates</th>
<th>N beginning</th>
<th>N end</th>
</tr>
</thead>
<tbody>
<tr>
<td>1660s–1670s</td>
<td>81</td>
<td>79 (100%)</td>
<td>58 (73%)</td>
<td>21 (27%)</td>
</tr>
<tr>
<td>1780s</td>
<td>100</td>
<td>95 (100%)</td>
<td>42 (44%)</td>
<td>53 (56%)</td>
</tr>
</tbody>
</table>

Figure 4. Location and date at the end of an eighteenth-century letter by Clara Sluijsken.
masc/fem) and desen “this” (acc sing masc), and abbreviations of anno Domini such as anno, ady, a, A°, which may also appear in front of the whole date or immediately preceding the year (10f–h).

(10) a. den 5 desember 1670  
    b. den 18 November 1780  
    c. 17 may 1665  
    d. den 25 may  
    e. 1672 den 1 november  
    f. desen 13 november A 1664  
       this 13 November anno 1664  
    g. adij den 12 Janu 1672  
       AD the 12 January 1672  
    h. Den 15 Januarij A° 1781  
       the 15 January anno 1781

Interestingly, a few examples appear where the original meaning of ady “anno Domini” must have been obscured by the abbreviated form, with reinterpretation as a preposition “on” (10i) or “in” as exemplified in (9e).

(10) i. Adij 10 Octobre Ano 1664  
       AD 10 October anno 1664  

(10i) is ambiguous on the assumption of repeated usage of “anno Domini” both in front of the month and the year. Such ambiguous examples may have been a bridging context between the original meaning and unambiguous examples of prepositional usage such as (9e).

**Additional information**

In very few cases, additional information is included in the date, once non-formulaic as in (11a) and twice formulaic such as the short formulaic expression in (11b).

(11) a. dit is onse 9 brief  
       this is our 9th letter  
    b. Met haest  
       with haste

6.1.3 Opening

The opening signals the start of the body of the letters, and more specifically the introductory part. As explained in Section 4, we consider opening formulae as well as closing formulae as text-type formulae which incorporate an intersubjective element through the link they establish with the intersubjective parts of the letter. Opening
formulae in the seventeenth-century subcorpus usually consist of three parts fulfilling the functions of addressing the recipient, greeting the recipient and referring to earlier communication. As described in Table 6 below, in the opening formulae of the eighteenth-century subcorpus, only two of these functions survive, which are moreover partly verbalised by other formulae. Reference to earlier communication is usually in the final position. Addressing and greeting the recipient can both come first. In the 81 letters in the seventeenth-century subcorpus, 52 begin with addressing the recipient, and 28 with greetings (one letter only has a reference to earlier communication). 37 letters contain both functions, and the majority of these (26) put the greeting formulae before the addressing formula. This section will discuss these three functions and the corresponding formulae.

Table 6. The functions of the opening of a letter in the seventeenth and eighteenth centuries.

<table>
<thead>
<tr>
<th>1660s-1670s</th>
<th>1780s</th>
</tr>
</thead>
<tbody>
<tr>
<td>address the recipient</td>
<td>address the recipient</td>
</tr>
<tr>
<td>greet the recipient</td>
<td>refer to (earlier) communication</td>
</tr>
<tr>
<td>refer to (earlier) communication</td>
<td></td>
</tr>
</tbody>
</table>

Addressing the recipient

Addressing the recipient proceeds in a fairly formalised way in the subcorpora used for the present study. The relevant formula comprises four positions: 1) the possessive pronoun mijn “my”, 2) one or more adjectives, 3) a noun indicating the function of the recipient or his relation to the writer, and 4) the recipient’s proper name. (12a) is an example taken from the seventeenth-century subcorpus, (12b) is from the eighteenth-century subcorpus.

(12) a. mijn lieue ende welb minde man jan van nes
    my dear and well-beloved husband Jan van Nes

b. Veel Geagte En lieve Vaader En Moeder
    very respected and dear father and mother

As can be seen in the examples, the adjectives may be modified by intensifying adverbs such as wel “very” and veel “very”. All elements of these formulae are also well-known from address and subscription formulae in early English correspondence (Nevala 2004b: 128–158).

Table 7 details the presence of the four positions in the addressing formulae found in the subcorpora used. Note that N adjective refers to the number of addressing formulae with at least one adjective, not to the total number of adjectives. In the seventeenth-century subcorpus, for example, no less than 97 adjectives appear in the 65
address formulae with at least one adjective, which obviously relates to the use of multiple adjectives as in (12a). As can be seen in Table 7, the possessive pronoun position and the proper name position sharply decrease in frequency in the eighteenth century, whereas the adjective and the function/relation position further increase to almost 100%. The number of positions, and thus the number of possible variants is reduced, resulting in a highly levelled formula in the eighteenth century.

The use of adjectives also seems to be patterned. There is a small set of adjectives that appear over and over again, the choice of which seems to be restricted by the function/relation. In Table 8, we present the adjectives in use in the seventeenth-century subcorpus, plotted against the type of relationship. Note that the seventeenth-century subcorpus used here comprises a small number of eleven business letters. There is a clear distribution of the adjectives used, ranging from neutral *eersame* “honest” to intimate *lieve* “sweet, dear”. The use of *goede* “good” and *goedgunstige* “gentle” seems to be restricted to relatively distant relations such as friends and business partners. Note that the concept of friendship was much more widely used in the Early Modern period, and did not imply the intimacy often associated with it in modern times (Kooijmans 1997).

### Table 7. Elements in the address formulae in the seventeenth and eighteenth centuries.

<table>
<thead>
<tr>
<th>Period</th>
<th>N addressing</th>
<th>N possess. pronoun</th>
<th>N adjective</th>
<th>N function/relation</th>
<th>N proper name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1660s-1670s</td>
<td>78</td>
<td>34</td>
<td>65</td>
<td>67</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>(100%)</td>
<td>(44%)</td>
<td>(83%)</td>
<td>(86%)</td>
<td>(73%)</td>
</tr>
<tr>
<td>1780s</td>
<td>100</td>
<td>1</td>
<td>98</td>
<td>98</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>(100%)</td>
<td>(1%)</td>
<td>(98%)</td>
<td>(98%)</td>
<td>(14%)</td>
</tr>
</tbody>
</table>

### Table 8. Adjectives used for different relationships between sender and recipient in the seventeenth century.

<table>
<thead>
<tr>
<th>Relationship</th>
<th><em>eersame</em> “honest”</th>
<th><em>waerde</em> “worthy”</th>
<th><em>beminde</em> “beloved”</th>
<th><em>lieve</em> “dear”</th>
<th><em>goede/goedgunstige</em> “good, gentle”</th>
</tr>
</thead>
<tbody>
<tr>
<td>spouses</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>family</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>friends</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>business</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Similar adjectives are found in the eighteenth-century subcorpus, with a similar distribution in accordance with the type of relationship; see Table 9. The eighteenth-century subcorpus does not contain any business letters. The main differences are that the neutral
adjective eersame “honest” is replaced by geagte “respected, esteemed”, while the relatively distant adjective in the eighteenth century is geerde “honoured”, not goede or goedgunstige. Also, lieve “dear” appears to have spread from being used between spouses to family relations in general. The subcorpora examined are fairly small, however, and primarily intended to demonstrate the formulaicity of epistolary language; too firm conclusions on diachronic changes should therefore not be drawn. Note, in this respect, that eersame “honest” does in fact appear in the eighteenth-century, though not in inside addresses but only in outside addresses.

**Greeting the recipient**

Greeting the recipient is a frequent function of seventeenth-century openings, while being completely absent from the openings of letters in the eighteenth-century subcorpus (see also Chapter 4, Section 3.2). In both periods, intersubjective greeting formulae appear further on in letters, which will be discussed in Section 7.2. Thus we need to distinguish between greetings as part of the text-constitutive opening formulae and greetings as part of the intersubjective formulae mostly found in the endings of letters. Apart from the different places these greetings occupy in the structure of letters, their different functions are also formally marked. The most frequent text-constitutive greeting in the openings of letters is presented in (13a), where the verbal elements sy gescheuen “be written” articulate the formula’s text-constitutive function (cf. Section 4 above). Intersubjective greeting formulae in the endings of letters lack this kind of meta-communicative explicitness (see Section 7.2).

(13) a. Een Vryendelijcke groetenysse sy gescheuen aen a friendly greeting be written to

Variants of this formula include modifiers such as seer “very” (thus “very friendly”) and different adjectives such as waerdeijge “worthy”. Its formulaicity also becomes clear from reduced variants that appear occasionally (13b, c).

(13) b. een vriendelicke groete nisse geschreuen an a friendly greeting written to
c. *een vriendelijke groeten an*
   a friendly greeting to

Greeting the recipient is an element in 39 out of 81 openings in the seventeenth-century subcorpus. Of these 39 text-constitutive greetings, 23 tokens (59%) are variants of the formula in (13a). The remaining 16 text-constitutive greetings are variants of two other formulae: *naer hertelijke groetenisse* “after cordial greeting” (9 instances) and the French loan *saluijt* (7 instances). Variants of the first formula include phrases with a different adjective (13d) or with the verb *wensen* “wish” (13e) or a combination of these (13f), reduced variants (13g), and one variant with a different preposition (13h), as well as the noun *wijensijnge* “(the) wishing”.

(13)  
   d. *naer vrindelijcke groetenis*  
      after cordial greeting  
   e. *Naer wensende alles Goedts*  
      after wishing all good  
   f. *naer groettenisse wensende*  
      after greeting wishing  
   g. *nae groetenis*  
      after greeting  
   h. *met wijensijnge alles goeds*  
      with wishing all good

If a letter contains both a formula for addressing the recipient and one for greeting him/her (37 tokens), the formula *Een Vryendelijcke groetenysse sy gescheuen aen* as in (13a) always precedes the addressing formula (23 tokens), whereas *saluijt* always follows it (6 tokens). The variants of *naer hertelijke groetenisse* may precede (4 tokens) as well as follow (4 tokens) the addressing formula.

*Referring to (earlier) communication*

A wide variety of formulae are used in the seventeenth-century subcorpus for referring to earlier communication. These include the formulae in (14a).

(14)  
   a. *Op 18 maij was aan U-E mijnen laasten daar aan mij refereere*  
      on 18 May was to you my last there to my referring  
      *sedert become U-E seer aengenamen vanden 28 october 1674*  
      since received your very pleasant of-the 28 October 1674  
      *daar uijt wij met groote alteratie gesien hebben*  
      there from we with great dismay seen have  
      ‘My last [letter] was to you on 18 May, to which I refer, since then [I have] received your very pleasant [letter] of 28 October 1674, in which we have seen with great dismay’
First of all, the object or objects of meta-communication, that is, the letter or letters previously exchanged, need to be characterised. This is commonly achieved by the use of a personal pronoun indicating the sender and an adjective which may be modified by an adverb. Note that the noun “letter” is absent. In (14a), the two most frequently used adjectives both appear, in the formulae mijnen laasten “my last” and U-E seer aengenamen “your very pleasant”, respectively. Variants of laasten include vorige “previous” and eerste “first”. The seventeenth-century subcorpus comprises 11 tokens of this formula.

Secondly, dates may be filled in – the date when the letter is received or the date when the letter was written. As can be seen in (14a), the preposition op “on” is used for the first date, the preposition van “from” for the second. The structure of the date following the preposition is mostly similar to the one outlined above (Section 6.1.2), viz. den-day-month-year, where in particular the article and the year may be left out. The seventeenth-century subcorpus includes 14 tokens of this formula.

Thirdly, some formulae explicitly state that the letter has been received. In (14a), sedert become “since then received” is an example. Other examples include soo daeteliijck Becoome “only just received”, wel terandt gecomen “come to hand well”, wel ter handt gestelt “lit. put to the hand well”, wel gewerde “come to hand well”, wel overghekomen “come over well”. There are 12 such phrases in the seventeenth-century subcorpus. In her selection of late-eighteenth-century letters, Austin (1973a: 15) found similar phrasings, which she refers to as ‘the formula “I received your (kind) letter of” + date’. Elspaß (2005: 165) mentions German examples from the nineteenth century (Deinen/Euren Brief, Dein/Euer Schreiben (vom...) habe ich/haben wir richtig erhalten/empfangen “your letter, your writing (of...) I/we have received well”).

Whereas these formulae are usually connected to the nounless formulae mentioned above, such as mijnen laasten “my last” and U-E seer aengenamen “your pleasant”, another formula which explicitly states that a letter has been received does incorporate the noun brief “letter”. An example is Ik heb u brief ontfangen “I have received your letter”, of which 5 tokens occur in the seventeenth-century subcorpus. Similarly, the writer may explicitly refer to the fact that he has previously sent a letter, see (14b).

(14) b. jck hebbe U.E myn lieue kint twee brieue gest
    I have you my dear child two letters sent
    ‘I have sent you, my dear child, two letters’

Again, the noun brieue “letters” is used. Five tokens of variants of this formula are found in the seventeenth-century subcorpus.

Fourthly, attention may need to be drawn to the contents of the previous letter(s), for which the formula daar uijt gesien “lit. there from seen, seen in it” as in (14a), and daer uijt verstaen “lit. there from understood, understood from it” are in use. The seventeenth-century subcorpus contains 9 variants of this formula. Elspaß (2005: 165) reports the use of und (...) (daraus/darin) gesehen/ersehen “and seen in/understood from it” in German nineteenth-century private letters.
Finally, in an act of meta-meta-communication, the formula *daar aan mij refereere* “lit. there to my referring, to which my reference, to which I refer” may be used to refer to the fact that earlier communication is referred to. Four tokens of this formula feature in the seventeenth-century subcorpus.

Many of the afore-mentioned formulae also appear in the eighteenth-century subcorpus; see examples (14c–h).

(14) c. *Myn laaste aan ue was*
   my last to you was

d. *Myn Voorige aan UEd: was*
   my previous to you was

e. *ik heb u eedele brief wel ontfangen*
   I have your letter well received

f. *Ik heb UEdlen brief van den 20 january 1780 wel ontfangen*
   I have your letter of the 20 January 1780 well received

g. *waar uijt wij zaage da*
   where from we saw that

h. *en daar uit te verstaan dat*
   and there from to understand that
   ‘and to understand from it’

The openings in eighteenth-century letters, however, differ fundamentally from the openings in seventeenth-century letters, in that greetings, which appeared in 39 out of 81 openings in the seventeenth-century subcorpus, are completely absent from openings in the eighteenth century. These greetings often established the transition from the text-constitutive parts (address, date) to the body of the letter, notably the introduction. This became apparent from phrases such as *sij gescheuen* “be written”, which draw attention to the fact that the greeting is part of a particular text, viz. the letter that the recipient is holding in his hands. In the eighteenth century, this transition needs to be established by other parts of the opening. As a result, reference to *earlier* communication is often replaced by reference to *current* communication, or put differently, by self-reference. This means that, after addressing the recipient, a formula follows that draws attention to the letter at hand. Five major types of formulae used for this function occur. Some of these occasionally appear in seventeenth-century letters (e.g. (14i)), albeit not in the subcorpus used for the present chapter. All of them feature prominently in the eighteenth-century letters in our corpus.

The most frequently used self-reference formula (17 tokens) in the eighteenth-century subcorpus is *deze dient* “this serves”, consisting of the pronoun *deze* and a form of the verb *dienen* (14i–k). Occasionally, a nominal derivation of *dienen* is used (14l) and reduced forms are also found (14m). The formula is mostly followed by a non-finite subclause introduced by *om* “to” or *tot* “to” (14i, j, l, m), but a finite subclause may also be used (14k).
(14) i. **dese dient om**
   this serves to

j. **Deze dienende tot**
   this serving to

k. **js deese dienende dat**
   is this serving that

l. **dit is dienstig om**
   this is useful to

m. **Dese om**
   this to

An alternative formula is found slightly less frequently (7 tokens), viz. **ik kan niet nalaten te schrijven** “I cannot refrain from writing” (14n, o), where sometimes the verb is replaced by **mankeere** “fail” as in (14p). The formula resembles the German nineteenth-century example **Ich kann/Wir können es nicht unterlassen, Dir/Euch wieder zu schreiben** “I/we can not refrain from writing to you again” (Elspaß 2005: 165).

(14) n. **ik kan niet naalaten om mijn lieve vrou te schrijven als tat**
   I can not refrain to my dear wife to write that
   ‘I can’t refrain from writing to my dear wife that’

o. **jk kan u niet Nalate om een letterje te schrijven hls dat**
   I can you not refrain to a letter to write that
   ‘I can’t refrain from writing you a letter that’

p. **ik kan niet mankeere met deese ocasie te schrijve als dat**
   I can not fail with this occasion to write that
   ‘I can’t fail, on this occasion, to write that’

Two more self-reference formulae which establish the transition from the text-constitutive part to the intersubjective part are in (14q, r): **neeme de vryheid te schrijven** “take the liberty to write” and **hebbe het genoegte melde** “have the pleasure to inform”. Of both formulae three tokens occur in the eighteenth-century subcorpus.

(14) q. **neeme ik de vryheid uwelEdele deese te schrijven**
   take I the liberty your honour this to write
   ‘Take the liberty Your Honour to write this’

r. **hebbe het genoegte UEde te melden, dat**
   have the pleasure you to inform that
   ‘[I] have the pleasure to inform you that’

The following and last self-reference formula, with no fewer than 15 tokens in the eighteenth-century subcorpus, is of a different kind. As will be explained below (Section 7.1), letters in the seventeenth-century subcorpus are characterised by
extensive health formulae (cf. lines 14–18 in the sample letter in (1) above). These elaborate passages on health virtually disappear in the eighteenth century (see Chapter 4). This does, however, not mean that the health theme is completely absent in eighteenth-century letters. Rather, it is inventively incorporated in opening phrases such as (14s). Fixed elements in these phrases are the pronoun deze “this” (cf. (14i–m)), which establishes the self-reference, and the wish or hope that the recipient will receive or find the letter in good health.

(14) s.  *Hoope ue dese ingesontheyt moogt ontfangen*
   hope you this in health may receive
   ‘[I] hope you may receive this in health’

   t.  *wenzen u Deeze en een goede welstan mag Aan treffen*
   wish you this in a good well-being may find
   ‘[I] wish you may find this while being in good health’

   u.  *Hoop & wense dat deese u myn waerde lief by goede gezonthyt, & all wenselyke wellzyn, mag in handen komen*
   hope and wish that this you my worthy love in good health and all desirable well-being may in hands come
   ‘[I] hope and wish that this may come into your hands, my worthy love, while you are in good health and all desirable well-being’

As the opening clauses of letters, these formulae fulfil a text-constitutive function. They typify the genre of letters, and draw attention to the fact that the document at hand is a letter. At the same time, by incorporating the health theme, they establish the transition to the intersubjective part, where the extensive health formulae typical of the seventeenth century are hardly in use any longer.

6.1.4 Closing
The canonical elements of the closings of letters are the actual closing formula and the signature. Where the distinction between the opening and the intersubjective part of the introduction is usually fairly easily drawn, the transition from the letters’ ending to the actual closing is more often than not difficult to discern. Recall example (1) at the beginning of the present chapter with its largely formulaic ending, which still incorporated non-formulaic context-dependent information. Furthermore, the endings often contain greetings and wishes – wishes that the recipient will continue to be healthy, wishes that the writer and the recipient will see each other soon again – larded with non-formulaic messages. For these reasons, we do not consider any greeting a closing formula. Instead, we consider most greetings to be intersubjective formulae (see 7.2). Only when they are clearly separated from the preceding discourse and immediately preceding the signature do we consider greetings as closing formulae. These, along with other closing formulae, will be discussed in the present section.
The eighteenth-century pattern does not differ substantially from the seventeenth-century convention for closing a letter. We will, therefore, first discuss in detail the closing formulae in the seventeenth-century subcorpus and subsequently deal with a particular eighteenth-century innovation.

In the seventeenth-century subcorpus, we find 78 closings. Closing formulae may consist of six elements: 1) one or more short idioms (see (15b) below), 2) a possessive pronoun, usually *ue* “your” or *ul* “your”, 3) one or more adjectives, which may be modified by adverbs, 4) a noun indicating the social role of the writer, that is his profession or his relation to the recipient, 5) the sender’s proper name, 6) the idiom *wat ick vermagh* “whatever I am capable of”. As before, resemblances to English seventeenth-century subscription formulae are obvious; for instance *thy affectionate and faithful husband/servant* and *your obedient and most truly affectionate daughter* (Nevala 2004b: 141). Elspaß (2005: 161, 163) provides similar examples from nineteenth-century German, such as *Euer Sohn und Bruder F. Dunker* “your son and brother F. Dunker”, and *Eure gehorsame Tochter Barbara Rueß* “your obedient daughter Barbara Rueß”.

The elements 2) to 5) are quite similar to those in use for addressing the recipient (see Sections 6.1.1 and 6.1.3). Obviously, second person pronouns (64 tokens) are used instead of first person pronouns. The range of adjectives used (48 tokens) is dominated by expressions indicating servitude and obedience. The most frequent adjectives are *dienstwillige* “willing, obedient” (14 tokens), *(toe)genegen(ste)* “willing, affectionate” (7 tokens) and *onderdanighe* “obedient, humble” (6 tokens). *Dienstwillige* in particular must have been so frequently used that a reduced variant *DW* came into existence, which was then often combined with *UE* “your” into one orthographical unit, e.g. *U.E.D.W.* (15a).

(15) a. *U.E.D.W. vrient [...] Michiel van Amelant*
   your obedient friend Michiel van Amelant

The nouns indicating social role (64 tokens) include *huijsvrou* “wife”, *man* “husband”, *broeder* “brother”, *frere* “brother”, *moder* “mother”, *kinderen* “children” and so on, as well as *dienaer* “servant”. The writer’s proper name appears in all 78 closings. The sixth possible element, the idiom *wat ick vermagh* “what I am capable of” or variants such as *wat wij vermogen tot uen besten* “whatever we are capable of for your good”, is found in only 7 closings. In the eighteenth-century subcorpus, not a single token occurs. This suggests that this formula, well-known from sixteenth- and early seventeenth-century usage, was already decreasing in the second half of the seventeenth century.

The first possible closing element, that of one or more short idioms, is realised 48 times in the seventeenth-century subcorpus. Five idioms are in use (15b) and the numbers indicate the frequencies.

(15) b. *bij mijn (16), (ver)blijve (16), hondert duijsent goede naght (6),*
   by me stay, remain hundred thousand good night
   *amen (6), vaert wel (4) *
   amen farewell, goodbye
Obviously, many of these are found throughout Europe, with, for instance, *I remain* (Austin 1973b: 131–132) and *verbleibe* (Elspaß 2005: 162) as counterparts of *(ver)* *blijve*. A full seventeenth-century closing is presented in (15c), while (15d) shows that the idioms of (15b) are sometimes combined.

(15) c. *By myn Jan Jansen van tiel ul vader wat ick vermach*  
*by me Jan Jansen van tiel your father whatever I can*

d. *hondert duijsent goede naght bij mij ul huysvrou sartie jans*  
*hundred thousand good night by me your wife Sartie Jans*

As stated above, the data from the eighteenth-century subcorpus do not differ substantially from the seventeenth-century data, apart from an important innovation: repetition of the addressing of the recipient. Immediately preceding the pronoun, therefore in the second position of the afore-mentioned closing pattern of the seventeenth-century subcorpus, the recipient is addressed in a way similar to the address in the opening of the letter (Section 6.1.3). Such repeated addresses appear in 50 out of 100 letters in the eighteenth-century subcorpus. See (15e) for an example; the arrow indicates the repeated address.

(15) e. *... gelijk ik meede de Eere heb mij*  
*... as I also the honour have me*

*met de opregste gevoelens van Eerbied en Estime*  
*with the sincerest feelings of respect and esteem*

*te noemen.*  
*to call*

→ *Mevrouw en Zeer G’eerde Vriendin!*  
*Madam and very honoured friend*

*uwelEdeles zeer Gehoorsaame*  
*your honour very obedient*

*Dienaar en vriend*  
*servant and friend*

*Jan Pieter Humbert.*  
*Jan Pieter Humbert*

‘as I also have the honour of calling myself, with the sincerest feelings of respect and esteem, Madam and very honoured friend!, your honour’s very obedient servant and friend Jan Pieter Humbert.’

As with the formulae in the previous section (e.g. (14i)), this new usage occasionally appears as early as in the seventeenth century, but its frequency rises to a remarkable level in the eighteenth century.
6.2 Text-structural formulae

A number of text-structural formulae mark the text structure by realising the transition from one part of the discourse to another. As stated above, these may in principle be used throughout the text and, in fact, in any text, but in practice, they are highly specific for letters and largely confined to specific text-structural functions. There are text-structural functions which initiate discourse (6.2.1), continue discourse (6.2.2), and end discourse (6.2.3). The first are mostly found in the introductions of letters, the latter in endings, and the second in between.

6.2.1 Initiating discourse

The most frequently used formula for initiating discourse runs *ick laet ul weeten als dat* “I let you know that”, see (16a), where the relevant excerpt from the sample letter in (1) is repeated. The formula is also found in both English and German letters of the Early and Late Modern period, for instance *to let you know*, *um dich/euch wissen zu lassen* “to let you know”, *thue ich euch ze wissen* “I let you know” (Austin 1973a: 16; Elspaß 2005: 165, 168–170), and also in Finnish letters from the nineteenth century (Laitinen & Nordlund 2012).

(16) a. *ick ul husvrou angenietge cornelis laet ul weeten als dat*
    I your wife Angenietge Cornelis let you know that

A total of 64 tokens of this formula were found in the 81 letters of the seventeenth-century subcorpus. Due to its relatively high frequency, this formula is suitable for sociolinguistic research; see Chapter 4, Section 3.6, where more examples are presented. Of the 64 tokens, 32 immediately follow the opening of the letter, confirming that the formula is often used in introductions, and thus for initiating discourse.

Nonetheless, 32 tokens are found in subsequent parts of the letters. In such cases, the formula often begins with the adverb *voorts* “further” or variants such as *voort, voorder, vorderen, verders* (16b), or with a form of this adverb plus expletive *soo* “so” (16c). In these cases, inversion of the subject and the finite verb is triggered by the adverbial phrase. When placed in subsequent parts of a letter, for instance in the middle part or even in the ending, the formula does not so much initiate discourse as such but rather introduces a new topic (see also Section 6.2.2). The same topic-shifting function is fulfilled by its English counterpart *to let you know*, as shown by Austin (1973a: 16) in her study of eighteenth-century letters.

(16) b. *voort laet ick v l weeten dat*
    further let I you know that

(16) c. *voort soo laet ick u l weeten als dat*
    further so let I you know that

There is a clear preference for the causative verb *laeten* “let”, though three alternative instances with the verb *doen* “do” are found in the seventeenth-century subcorpus (16d).
Chapter 3. Epistolary formulae

(16) d. ick maritien Jans u l huusvrou doe u hier mede te weten
I Maritien Jans your wife do you here with to know
‘I, Maritien Jans, your wife, let you hereby know’

The formula is still fairly frequent in the eighteenth-century subcorpus, although, with 34 tokens in 100 eighteenth-century letters, it is not as pervasive as in the seventeenth-century subcorpus. The frequency difference will be discussed in detail in Chapter 4, Section 3.6. There are no formal differences between the seventeenth-century variants and the eighteenth-century variants, except for orthographical changes, but a remarkable usage should be mentioned here. In Section 6.1.3, we argued that the text-constitutive self-reference formula deze dient came to be used quite often in the eighteenth century; cf. examples (14i–m). This self-reference formula is found combined with the let know formula in a non-finite subclause (16e, 16f), thus marking the transition from self-reference (‘this is a letter’) to the actual intersubjective contents of the letter.

(16) e. dese diend om UEd te Laten Weten
this serves to you to let know
‘this serves to let you know’

f. dese dient om ue te doen weeten als dat
this serves to you to do know that
‘this serves to let you know that’

6.2.2 Continuing discourse

As explained in the previous section (examples (16b, c)), discourse continuity may be established by the concise formulaic expression voorts “further” or its variants, sometimes combined with expletive soo “so”. The use of this formula signals to the reader that a new topic is about to be introduced. The two main variants are voors (soo) “further (so)” and verders “further”, the first dominating the seventeenth-century data, the latter the eighteenth-century data. Like the I let you know that formula (Section 6.2.1), the formula voorts (soo)/verders is similarly highly frequent and therefore suitable for sociolinguistic research (see Chapter 4, Section 3.7). A second formula establishing discourse continuity by introducing a new topic may be translated as “as far as I am concerned” or “concerning”.

In the seventeenth-century subcorpus, as many as 90 tokens of the voors (soo)-formula occur. Of these 47 are followed by non-formulaic clauses conveying context-dependent information and confirming the formula’s usefulness for continuing discourse while switching to a new topic; see (17a,b).

(17) a. voorders soo zynder al veel van kennesse gestorven
further so are-there already many of acquaintance died
‘furthermore, many acquaintances have already passed away’

b. vorders stiere ick vl twee kinneties harincmet twee tonneties
further send I you two kilderkins herring-with two barrels
boter
butter
‘furthermore, I send you two kilderkins of herring with two barrels of butter’

At the same time, the formula is subject to strong collocational patterns in that 43 out of 90 tokens (48%) are followed by another formula, viz. in 25 cases by *I let you know that* (see 6.2.1, examples (16b, c)), recalling German nineteenth-century formulae such as *ferner tue ich euch zu wissen* “further I let you know that” (Elspaß 2005: 168). In 18 cases, it is followed by a formula with the keyword *schrijven* “write, writing”. The latter formula comes in various subtypes. The two main subtypes are in (17c) and (17d), where the second is a negative alternative to the first. Furthermore, reduced variants (17e, f) illustrate the formula’s high frequency.

(17) c. voort ijs mijn schrijven als dat
   further is my writing that

d. vorder en wet ick niet vel mer te scrijven dan
   further NEG know I not much more to write than
   ‘furthermore, I do not know much more to write than’

e. Voorder niet meer
   further not more

f. vorders niet als
   further nothing than

Examples such as (17d–f) bear a striking resemblance to formulae used for ending discourse (see 6.2.3).

In the eighteenth-century subcorpus, 46 tokens of the *verders* formula were found. No fewer than 42 of these (91%) are followed by another formula; this is much more frequent than in the seventeenth century, suggesting that the use of this lexeme has become even more formulaic over time. As in the previous century, *verders* may be followed by either a new piece of information ((17g), cf. (17c)) or by the statement that there is no new information to give ((17h), cf. (17d–f)). These two major subtypes are responsible for 19 tokens in total.

(17) g. Verder heb ik u, myn hertie lief te melden, dat
   further have I you my heart love to inform that
   ‘furthermore, I must inform my dear love that’

h. Verders niets sonderling meer te melden hebbende als
   further nothing special more to inform having than
   ‘Furthermore, I have nothing special to say but’

Furthermore, the *verders* formula appears in combination with formulae intended to end the discourse (17i), which will be discussed in the next section.
(17) i. *verders Breek ik aff*  
  further break I off  
  ‘I furthermore end [this letter]’

Finally, *verders* may introduce intersubjective formulae such as greetings (17j) and health formulae (17k); see also Section 7.

(17) j. *Verders Zyt duyzend Maal gegroet En gekust*  
  further be thousand times greeted and kissed

k. *Verders zijn alle nog Gesond en Wel*  
  further are all still healthy and well

The other formula used for continuing discourse, namely “as far as I am concerned” or “concerning”, features strongly in the eighteenth-century subcorpus where we find 29 examples. Again, two major subtypes can be distinguished. The first subtype (19 tokens) contains the phrase *wat X aangaat* “what X concerns”, where X usually is or contains a pronoun of the first person (17l). The second subtype (10 tokens) comprises the originally present participle of the same verb *aangaan* “concern”, which is however grammaticalised or grammaticalising into a preposition (17m, n). Again, it may be connected to a first person pronoun (17m), but in many cases it appears with noun phrases (17n).

(17) l. *wat ons en onze dogter Benevens mijn zoon Pouwlus aangaat*  
  what us and our daughter next to my son Pouwlus concerns

m. *myn aangaande*  
  me concerning

n. *aangaande de nalaatenschap van moeder*  
  concerning the inheritance of mother

6.2.3 Ending discourse

Just as there are epistolary formulae for taking up discourse (6.2.1) and for continuing discourse (6.2.2), likewise there are formulae for ending the discourse. A well-known seventeenth-century example, which also occurs in the sample letter in (1), is presented in (18a). A similar phrase in eighteenth-century English ran *so no more at present* (Austin 1973b: 129–130).

(18) a. *nijet mer op dijt pas dan*  
  not more on this step than  
  ‘nothing more for now, on this occasion than’

Variants include the use of *geen* “no” instead of *niet* or *nijet* “not”, and of *als* “as” instead of *dan* “than”, in line with the still existing *als/dan* variation after comparatives. The formula signals the end of the intersubjective part of the letter and announces the coming text-constitutive part, notably the closing formulae. But before the actual
closing formula appears, the ending formula in (18a) enables the writer to give one final piece of intersubjective information. Of the 21 tokens of this formula in the seventeenth-century subcorpus, 19 are followed by either the so-called commendation formula ((18b), cf. Section 8) or a greeting formula ((18c), cf. Section 7.2).

(18) b. *Geen meer op dit past den zijt de heere bevoolen*  
no more on this step than be the Lord commended  
‘nothing more for now than be commended into the hands of the Lord’

c. *nijet meer dan van harte gegroet*  
not more than from heart greeted  
‘nothing more than be greeted from the heart’

It is far from uncommon that after this formula and its collocational continuation as in (18b, c) the writer does not close the letter at all, but instead continues writing, and even gives new non-formulaic context-dependent information. A telling example is the sample letter in (1), where lines 32–34 indeed convey new information, only after which the writer closes the letter. As can be seen in (18c) and (18d), reduced variants are also used, which may however also be considered as reduced variants of (17d) in the previous section, cf. (17d, e).

(18) d. *niet meer als*  
not more than

The formula *niet meer op dit pas* “nothing more for now” no longer occurs in the eighteenth-century subcorpus. Instead, we find a few examples of *niet meer te schrijven* “nothing more to write”, which resembles formulae discussed in the previous section, e.g. (17h), but lacks the adverb *verders*.

(18) e. *Ik weet nu niet meer te schryfen*  
I know now not more to write

Other formulae, however, are much more salient in the eighteenth-century subcorpus, while hardly appearing in seventeenth-century letters. Firstly, there is a formula with the keyword *afbreken* “break off” (15 tokens). The most common variant is in (18f), but alternatives are also found (18g, h).

(18) f. *hier meede breek ik af*  
here with break I off  
‘with this I end’

g. *hiermede moet afbreken*  
here-with must break off  
‘with this [I] must end’

h. *hier mede af Breekende*  
here with off breaking  
‘ending with this’
Secondly, there is a formula with the keyword *eindigen* “end” (6 tokens), see (18i, j). This recalls the use of *I must conclude* in English eighteenth-century letters (Austin 1973b: 130), and the use of *für jetzt will ich schließen* “for now I want to close”, *so wil ich schlüsschen* and similar phrasings in nineteenth-century German letters (Elspaß 2005: 166).

(18) i. *Intussen zal ik Eijndige*  
meanwhile shall I end

j. *Edog Ik zal hiermede eindigen*  
but well I shall here-with end

Finally, there is a fairly frequent and complex formula (30 tokens), which is introduced by the preposition *na* or *naer* “after” or the subordinator *nadat* “after”, which is usually continued by a form of the auxiliary *hebben* “have”, and then rounded off by a form of either the verb *blijven* “remain” or the verb *noemen* “name, call”; see the examples in (18k, l).

(18) k. *naa uedle*  
after you
*des heeren segen te hebben toe gewenst*  
the Lord’s blessing to have wished
*blijve ik met alle agtin*  
remain I with all respect
‘after having wished for you the Lord’s blessing, I remain with all respect [repeated addressing, signature]’

l. *na dat ue duysen maal in my dedagte [sic] onthelst*  
after that you thousand times in my thoughts embraced
*hebben hier nevens Meen ik de my te noemen met alle agten*  
have here next think I ? me to call with all respect
*en eerbiet*  
and esteem
‘after having you embraced a thousand times in my thoughts, next to this I think I may call myself with all respect and esteem [repeated addressing, signature]’

The *na-hebben* part usually contains an intersubjective phrase, conveying a blessing (18k) or a greeting (18l). The forms of *blijven* (cf. English *remain*, German *verbleiben*) and *noemen* create the transition to the writer’s name and/or signature, in the eighteenth century commonly preceded by the repeated addressing of the recipient. Note that this complex ending formula is intertwined with phrases, especially forms of *blijven*, which already appeared in closing formulae in the seventeenth-century sub-corpus, see Section 6.1.4, example (15b). This means that these concise seventeenth-century closing formulae, which stood more or less apart from the preceding discourse, are extended into the ending part of letters, creating a more gradual transition to the closing of the letter.
As with many of the formulae discussed in the previous sections, reduced variants of this complex formula occur as well (18m).

\[(18) \text{ m. } na\text{ mij in UEdg altoos aan-}\]
\[\text{after me in your always en-}\]
\[\text{houdende vriendschap te hebben aen bevolen}\]
\[\text{during friendship to have commended}\]
\[\text{‘after having myself commended into your always enduring friendship}\]
\[\text{[repeated addressing, signature]}\]

In (18m), the rounding off element, viz. a form of blijven or of noemen, is missing.

7. **Intersubjective formulae**

In this section, we will discuss the three types of intersubjective formula found in our letter corpora, viz. health formulae (7.1), greeting formulae (7.2) and contact formulae (7.3).

7.1 **Health formulae**

The health theme is very prominent in the letters in our corpora, in particular in letters from the seventeenth century. The strong differences in frequency between the seventeenth and the eighteenth centuries will be discussed elsewhere; see Section 7.1.4 below and Chapter 4. Because of these diachronic differences, we will focus on seventeenth-century letters in Sections 7.1.1, 7.1.2 and 7.1.3.

After the opening of the letter, the most commonly broached subject is health, and the usual way of verbalising the subject is through formulaic language. Health formulae, which are also found in English (Davis 1965; Austin 1973a, 2004; Nevalainen 2001), consist of health statements and health wishes. These will be discussed in 7.1.1 and 7.1.2, respectively. But health formulae may also appear in the endings of letters. This means that we encounter not only two different types of health formulae (statements and wishes) but also two fixed places where health formulae typically appear: the introduction and the ending. We found 124 health formulae in the 81 letters from the seventeenth century used for the present chapter. In Table 10, we give the distribution of these 124 formulae according to the two variables type and place. As can be seen in Table 10, health statements are predominantly found in the introductions of letters (58/65) and health wishes mainly in letter endings (39/59).

7.1.1 **Health statements**

Health statements communicate the health of the writer (19a), of the writer and others (19b), solely of others (19c, d), or they may confirm the addressee’s health (19e). The
most commonly used expression for the state of health is the almost tautological formula *kloeck en gesont* “strong and healthy” (19a, b), though both elements may also occur separately (19c,d). It has a German nineteenth-century counterpart in the equally more or less tautological formula *gesund und munter* “healthy and lively”, with variants such as *gesund und wohl* “healthy and well” (Elspaß 2005: 165).

(19) a. *als dat jck noch klock en Ghesondt ben tot noch toe*  
that I still strong and healthy am until now

b. *als dat ick en u moeder en u vrou en ul soon en*  
that I and your mother and your wife and your son and

*alle goede vriende noch kloeck en gesont sijn*  
all good friends still strong and healthy are

c. *onse martije reijnders met de kijnderen sijn noch godt danck*  
our Martije Reijnders with the children are still God thank

*gesont*  
healthy

d. *moeder en susters en broeder en al onse urienda die sin noch*  
mother and sisters and brother and all our friends they are still

*klock*  
strong

e. *en daer uijt verstaen U:L: gesontheijjt*  
and there from understood your health

‘and [I have] understood your health [that you’re healthy] from it’

Of the 65 health statements found in the seventeenth-century letters (Table 10), no fewer than 43 use the formula *kloeck en gesont* (with some of these 43 containing only one of the two adjectives). Health statements often immediately follow a formula used for taking up discourse (Section 6.2.1), and they therefore mostly have the form of a subclause (19a, b), but main clauses may also be used (19c, d). For more examples and sociolinguistic analyses of the formula, we refer the reader to Chapter 4.

The other 22 health statements (out of a total of 65) are variants of less frequent formulae, of which we will discuss the three most important ones. The first one, *wel te passe* “lit. well to state, in good health” (19f), functions in a very similar way to *kloeck en gesont* in that it usually immediately follows a formula for initiating discourse.
(19) f. *dat jck noch wel te passe ben*
   that I still well to state am
   ‘that I am still in good health’

Another way of expressing health is by making use of the noun *gesontheijt* “health” in a collocation with the verb *verstaen* “understand” (19g). This collocation is often used when the addressee’s health, communicated in an earlier letter, needs to be confirmed (19e), but it may also occur when the writer communicates his own health (19g).

(19) g. *soo doen ijcck ul mijn gesoent heijt verstaen*
   thus do I you my health understand
   ‘thus I let you understand my health’

Finally, we wish to draw attention to examples where the wording is similar to previous examples, with the noun *gesontheijt* featuring prominently, but where in addition God is thanked for the state of health (19h, i).

(19) h. *van mij en mij kint en vors vaders en susters en broders*
   of me and my child and further father’s and sister’s and brother’s
   *gesontheijt hebben wij godt te dancken*
   health have we God to thank
   ‘we have to thank God for my and my child’s and furthermore for father’s
   and sister’s and brother’s health’

i. *voor mij en ons kint dank ijcck godt voor gesontheijt*
   for me and our child thank I God for health
   ‘I thank God for the health of myself and of our child’

Phrases such as (19h, i) are regarded as health formulae because the health theme appears to be dominant, but they also contain the Christian-ritual function. As explained in Section 4, formulae may combine the three main functions of epistolary formulae that we distinguished there (see also Section 8 below). Deciding which function prevails depends on the interpretation made. In cases such as (19h, i), the main message appears to be that everyone is still in good health, for which God is routinely thanked. Importantly, these phrases occur in the introductions to letters, immediately following a health wish to the addressee (19j, k), in other words, in a context which is predominantly focused on the health theme in almost any of the seventeenth-century letters.

(19) j. *u gesontheijt ben ick altit verhopende [(19h): van mij en mij kint en vors etc.]*
   your health am I always hoping
   ‘I am always hoping for your health [(19h): we have to thank God etc.]’

k. *ijck verhoope ul gesontheijt [(19i): voor mij en ons kint etc.]*
   I hope your health
   ‘I hope you are healthy, wish you good health [(19i: I thank God for the
   health of etc.)]’
Health wishes such as the first parts of (19j, k) are the topic of the next section.

7.1.2 Health wishes
The seventeenth-century letters that we explored for the present chapter contain a wide variety of health wishes. The 59 health wishes found can be arranged into four groups of formulae with similar lexical material and functions. The first group consists of formulae with either the verb *wensen* “wish” (20a) or *(ver)hopen* “hope, wish for” (20b), plus *gesont* “healthy” or *gesontheijt* “health” (20a, b). Sometimes the Frenchified form *dispositie* “disposition, state” is used, preceded by the adjective *goede* “good” (20c).

(20) a.  *wensse ve gesontheijjt*
    wish you health

b.  *verhopen ul gesontheijjt*
    hope, wish you health

c.  *verhoope evenwel ve goede dispositie*
    hope, wish still your good state
    ‘[I] still wish you a good disposition’

The second group of examples is very similar, albeit that it is not the addressee’s health in general that is wished for, but specifically his health while en route. The same lexical material applies, extended with verbs of movement such as *komen* “come” (20d) and *arriveren* “arrive” (20e).

(20) d.  *ick hoep dat ghi met gesontheit weeder thus komen mogen*
    I hope that you with health again home come may
    ‘I hope that you may come home again with health’

e.  *jck verhoeppe dat ul met voorspoet ende gesundtheydt tot midelburch sult gearriuerdt syn*
    I hope that you with prosperity and health at Middelburg will arrived be
    ‘I hope that you will have arrived at Middelburg with prosperity and health’

The third group still makes use of similar lexical items, but focuses on the reunion of the writer and the addressee, using collocations such as *sien en spreeken* “see and speak” (20f).

(20) f.  *dat wij malkan dere met gesontheijjt moge zien en spreeken*
    that we each other with health may see and speak
    ‘that we may see and speak to each other with health’

Examples such as (20f) will also be discussed in 7.3, where we will consider them as contact formulae. In such contact formulae, which express the wish that the writer and the addressee will see each other again, the health theme is incorporated by the
adverbial phrase *met gesontheijt* “with health”. Thus again we have an example of one formulaic expression fulfilling different functions (cf. Section 4).

Such a combination of functions is also found in the fourth group of health wishes, where the health theme is combined with the Christian-ritual function and God is invoked. There is a syntactic shift in that God is mostly presented as the subject of the clause. The wish that God will look after the addressee is expressed by either a subjunctive form of *geven* “give”, *verlenen* “grant” or *sparen* “spare, save” (20g), or a modal verb such as *zullen* “shall” or *wollen* “want” plus the infinitive of one of these verbs (20h, i, j). There is also a lexical shift in that *bewaring* “keeping” and *bewaren* “keep” are often used instead of phrases with *gesontheijt* “health” (20j).

(20) g. *die* [= God] *spaer* ue *mijn lieve man* in *langhdurende* he save-subj you my dear husband in enduring

  *gesontheijt* health

‘may he save you, my dear husband, in enduring health’

h. *die* [= God] *ick hoop vl* *gesontheit en behouwe reis* *sal geeue* who I hope you health and safe journey will give

‘who, I hope, will give you health and a save journey’

i. *ijck hoep dat dije goede godt v e* *sal gesontheijt verlen* I hope that the good God you will health grant

‘I hope that the good God will grant you health’

j. *godt ope ope sal u beijde beware* God hope hope will you both keep

‘God, [I] hope, [I] hope, will keep you both’

7.1.3 *Subordinate health formulae*

Health formulae are characteristic of seventeenth-century letters in particular and this also applies to the so-called *subordinate* health formulae (cf. Kuiper 2009: 7). To gain an idea of what we mean by subordinate health formulae, consider (21), where we present the first part of the introduction of two different letters. Both letters contain formulae relating to the health theme. These formulae are successively represented in italics, in boldface, underlined, in small caps and again in italics. Translations closely following the Dutch original are given below with the same layout. The phrases in normal font are also formulaic, but they are not health formulae and will be disregarded here (but see Sections 6.1.3 on letters’ openings and 6.2.1 on initiating discourse).5

---

5. The last sentence of Example 2 (“And I hope that the good God will make this last for a long time”) looks very much like a subordinate health formula. However, we have not found any other token of this supposed formula.
Chapter 3. Epistolary formulae

(21) The first part of the introduction of two seventeenth-century letters

**Example 1:** letter from Kathelijnen Haeswants, 10 November 1664
Vriendelijke Groetijssse aen ul mijn lieve ende bemijnde man Leendert arijensen haeswant jck katelijinghen haeswants ul hujsvrou late ul weten als dat jck met al onse kijnderen noch kloock ben godt lof van sijn genade ver-hoepende dat het met ul oock soo is ware het anders ’t soude mijn van harten leet sijn om hoeren dat weet godt almachtijch die een kender van alle harten js

**Translation 1**
A friendly greeting to you, my dear and beloved husband Leendert Arijensen Haeswant. I, Kathelijghen Haeswants, your wife, let you know that I with all our children am still in good health. Praise the Lord for his mercy. Hoping that you are also like that. If it were different, I would be very sorry to hear that. The almighty God, who knows all the hearts, knows this.

**Example 2:** letter from Angenietge Cornelis, 15 September 1664
eersaemen seer beminden man roellant iosten oost voren dick en alderliesten man ick ul husvrou angenietge cornelis laet ul weeten als dat ick en ul vaeder en min vaeder en moeder noch klock en gesont sien en ick hoep min alderliesten man roellant iosten oost vooren dick oock mede soo is waer het anders het souden min niet lief om hooren weesen en ick hoep dat het den goeden goed godt langen sal laeten dueren

**Translation 2**
Honest and very beloved husband Roellant Iosten Oostvoorendick and most dearest husband. I, your wife Angenietge Cornelis, let you know that I and your father and my father and mother are still in good health. I hope my dearest husband Roellant Iosten Oostvoorendick is also like that. If it were different, I would regret having to hear that. And I hope that the good God will make this last for a long time.

As becomes clear from (21), we can distinguish between health formulae and subordinate health formulae. Only the first phrase, in italics, contains an overt reference to the health theme, with the common use of the words kloock “strong” and gesont “healthy”.

Figure 5. First part of a letter by Kathelijne Haexwant, dated 10 November 1664.
The subordinate formulae do not contain an overt reference to the health theme, and are merely related to this theme by their reference to the first formula, where the health theme is made explicit. As such, they index a specific discourse structure, consisting of “sequential formulaic dependencies” (Kuiper 2009: 7). This means that the introductory passages of letters are not formulaic simply in that they use similar language, but also in that they constitute a formulaic genre, with discourse structure rules governing the sequence of formulae (Kuiper 2009: 7).

Focusing on these sequences in introductory passages, strong formulaic patterns can be found. Not every letter contains all the subordinate health formulae, but many letters contain some of them. The maximum length of an introductory formulaic health passage is six formulae, see Table 11, where the frequency of the subordinate formulae in the 81 seventeenth-century letters is presented. As with other formulae, there is some variation in wording, the examples in Table 11 representing the most frequent variants. The formulae 2–6 are almost only found immediately following an obvious health formula such as formula 1 in Table 11. Note that formulae 2 and 6 could also be considered as Christian-ritual formulae, but since they only appear in this specific discourse structure, we prefer to consider them as subordinate health formulae instead. Similar formulae are found in English letters from the fifteenth to the early nineteenth century (Davis 1965; Nevalainen 2001; Austin 1973a, 1973b, 2004). Laitinen & Nordlund (2012) provide examples from nineteenth-century Finnish. Elspaß (2005) does not mention German counterparts from his nineteenth-century corpus.

Table 11. (Subordinate) Health formulae in introductory passages in the seventeenth-century.

<table>
<thead>
<tr>
<th>Sequence of (subordinate) health formulae in introductory passages</th>
<th>Subordinate health formulae:</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>als dat ick en ul vaeder en min vaeder en moeder noch klock en gesont sein</td>
</tr>
<tr>
<td></td>
<td>‘that I and your father and my father and mother are still in good health’</td>
</tr>
<tr>
<td>2</td>
<td>godt sij lof van syn groote genade</td>
</tr>
<tr>
<td></td>
<td>‘Praise the Lord for his mercy’</td>
</tr>
<tr>
<td>3</td>
<td>gelijck ik hoop van vl mijn lief te verstaen</td>
</tr>
<tr>
<td></td>
<td>‘as I hope to hear from you my love’</td>
</tr>
<tr>
<td>4</td>
<td>het welcke mijn van harten seer lief om te hooren is</td>
</tr>
<tr>
<td></td>
<td>‘which I [would] very much love to hear’</td>
</tr>
<tr>
<td>5</td>
<td>waer het Anders het waer ons van herten leet</td>
</tr>
<tr>
<td></td>
<td>‘if it were different, we would be very sorry’</td>
</tr>
<tr>
<td>6</td>
<td>dat wedt godt almachtich die en kender van alle harten is</td>
</tr>
<tr>
<td></td>
<td>‘The almighty God, who knows all the hearts, knows this’</td>
</tr>
</tbody>
</table>

Total subordinate formulae 95
In total, no fewer than 95 of these subordinate health formulae were found in the seventeenth-century letters explored for the present chapter. Note that these are not included in the 124 health formulae found in these 81 letters, and discussed in the previous sections (cf. also Table 10). Put differently, in this corpus 219 formulae occur either directly or indirectly relating to the health theme.

7.1.4 Health formulae in the eighteenth century
One of the most salient differences in formulaic language between the seventeenth and the eighteenth centuries is the relative lack of health formulae in the eighteenth century. Of all formulae discussed in the previous sections, only a handful remain in the eighteenth century. Still, as explained in Section 6.1.3, the health theme is not completely absent from eighteenth-century usage. At the end of that section, we drew attention to formulae such as (22a), where the health theme is incorporated into the opening phrase.

(22) a. *Hoope ue dese ingesontheyt moogt ontfangen*
    hope you this in-health may receive

Another remarkable difference is a change in the expression *kloeck en gesont* “strong and healthy”. This tautological formula, commonly used in the seventeenth century (Section 7.1.1), does occur in the eighteenth century, but the first adjective is replaced by *fris* “fresh”; see (22b, c).

(22) b. *den 3e Maij deezen alhier fris en gezond aangeland ben*
    the 3rd May this here fresh and healthy arrived am
    ‘[I] have arrived here (being) fresh and healthy on May 3rd’

(22) c. *dese diend om UEd te Laten Weten dat ik Nog fris en gezond zyn*
    this serves to you to let know that I still fresh and healthy am
    ‘this [letter] serves to let you know that I am still fresh and healthy’

The new collocation *fris en gezond* “fresh and healthy” will be discussed in Chapter 4, where we will also deal with diachronic differences of formulaic language.

7.2 Greeting formulae
As described in Sections 6.1.3 and 6.1.4, greeting formulae frequently occur in the openings and closings of letters. In addition, greeting formulae are used throughout letters, and notably in the third main part, the ending. Apart from the greeting formulae found in openings and closings, no less than 116 greeting formulae appear in the 81 seventeenth-century letters, often accumulating towards the end of the letter. Unsurprisingly, many different phrases are used for greeting the addressee and/or others, but
at the same time, highly conventionalized patterns arise. In this section, we will present illustrative examples of the main types of greeting formula. We will primarily use seventeenth-century examples, stressing that the eighteenth-century examples would be very similar. The basic formulae used for greeting did not change in the course of time.

The first set of formulaic phrases (12 tokens) consists of imperatives of the verb *sijn* “be” plus the participle *gegroet* “greeted” (23a), where the participle is mostly modified by *van harten* “from the heart” (23a) or *hertelijk* “heartily” (23b), sometimes by *vrindelick* “friendly” (23c). Reduced forms lacking the imperative (23d) also frequently appear (8 tokens).

(23) a. *sijt van harten gegroet*
   be from heart greeted

b. *soo sijt hier mede seer hertelijck gegroet*
   so be here with very heartily greeted

c. *sijet van mij vrindelick gegroet*
   be from me friendly greeted

d. *van harten gegroet en duisent mael gekust*
   from heart greeted and thousand times kissed

We also find the simple phrase the greeting(s) to/from (8 tokens) as in (23e). Variants include clauses with the same noun *groetenisse* “greeting” and a form of the verb *doen* “do” (23f), and with the verb *groeten* “greet” (23g).

(23) e. *de groetenes van ul hueijsurou en cenders en ael vrenden*
   the greeting from your wife and children and all friends

f. *doedt de groetenisse van min moeder aen min vader en aen uwer liede en van ons allen en aen alle goe vrienden*
   do the greeting of my mother to my father and to you people and from us all and to all good friends

g. *groet onse kernels en al frijnde en bekende*
   greet our Kernels [= Cornelis] and all friends and acquaintances

In another group of related constructions (15 tokens), the verbs used are *wensen* “wish” (23h, j) and *seggen* “say” (23k, l). That which is wished for or said typically runs *hondert duissent goede nacht* “one hundred thousand good nights” (23h, i) or *veel goede nacht* “many good nights” (23j, k), sometimes simply *goeden nach* “good night” (23l). Reduced forms also appear (23i), and on the other hand extensive greetings may be communicated, for instance by adding a long series of names (23l).

(23) h. *wenssen Ul hondert duissent goede nacht*
   wish you hundred thousand good night

i. *hondert duissant goede nacht*
   hundred thousand good night
Chapter 3. Epistolary formulae

j. mijn suster marrijtien en haer man corneelis wenschen v l
   my sister Marrijtien and her husband Corneelis wish you
   veel goede nacht
   many good night

k. segt u maet en alle goede kennis veel goede nacht
   say your mate and all good acquaintances many good night

l. seght vader ende moede goeden nach ende mijn broeder Jeroen
   say father and mother good night and my brother Jeroen
   ende ul broer Jacob schelwaert met sijn vrouwe ende
   and your brother Jacob Schelwaert with his wife and
   kees de vrijese met sijn hus vrouwe ende moeder de ploe
   Kees de Vrijese with his wife and mother De Ploe
   in mechellen ende moeij Janneke koster ende ul moeije karstie
   in Mechelen and aunt Janneke Koster and your aunt Karstie
   ende alle goede vrienden
   and all good friends

Occasionally, greetings are conveyed by using the verb soenen “kiss” (23m, n).

(23) m. ick sal u soon eens soene
   I will your son once kiss
   ‘I’ll kiss your son’

   n. ijck sal onse kijndt eens soenen voor u
   I will our child once kiss for you
   ‘I’ll kiss our child in your name’

Finally, seemingly common phrases such as adieu “adieu” may occur, variants of which are still found in many present-day dialects (e.g. adië, aju, ajuus), but it is shown that these phrases, with just six instances in the seventeenth-century letters, were less commonly used in written language than the afore-mentioned greeting formulae. In this respect, it is interesting that in the 100 eighteenth-century letters, variants of zijt ge-groet “be greeted” as in (23a) constitute quite a large category with no fewer than 21 tokens. With its subjunctive verb form, this formula seems characteristic of written language, rather than spoken language. It appears that even in the intersubjective domain of greeting, which must have prominently featured in both written and spoken language, conventions of the written language outweighed conversational routines.

7.3 Contact formulae

Apart from health and greetings, the wish for contact is the third main intersubjective theme addressed in the letters in our corpus. And as with health and greetings, a vast body of formulae were in use. In this section, we will discuss these contact formulae,
mainly drawing on the seventeenth-century subcorpus, but where appropriate discussing eighteenth-century examples as well.

The shortest contact formula is probably *schrijf eens* “lit. write once, please write”, expressing the wish that the addressee will write a letter to the sender of the present letter. It resembles German phrases such as *schreibe mir aber gleich Antwort* “lit. write me but immediately answer, please send me an answer as soon as possible”, found in nineteenth-century private letters (Elspaß 2005: 167). The wish can be specified by an indication of the topic of the writing, as in the following example from our seventeenth-century subcorpus (24a).

(24) a. *schrijf eens hoe dat ghij lieden ghevaren zijt met daniel
   Tissen
   ‘please write how you have traveled/got on with Daniel Tissen’*

Variants of this formula appear seven times in the 81 seventeenth-century letters and eight times in the 100 letters from the eighteenth century. More frequent occurrences are variants of a formula that usually contains at least the elements *dat wij malkanderen* “that we each other”. Two basic subtypes of this formula feature, the first expressing the wish that the writer and the sender will see (and speak to) each other again (24b), the second conveying the wish that they will simply be together again (24c).

(24) b. *dat wij malkanderen sullen zien en spreken
   that we each other will see and speak

   c. *dat wi met gesontheit bi malkanderen sullen komen
   that we with health to each other will come

Variants of the formula appear 24 times in the 81 letters from the seventeenth century, and 19 times in the 100 letters from the eighteenth century. As a subclause, the formula needs to be introduced; it is most often preceded by either *ick hoop* “I hope” or by *got geve* “God give+SUBJ, may God give”. When the latter phrase is used, the Christian-ritual theme (see Section 8 and the next paragraph) is connected to the contact theme.

The formula may furthermore be combined with elements referring to health, thereby blending two intersubjective domains. In its most simple and most common form, this reference consists of the adverbial phrase *met gesontheit* “in health” (24d).

(24) d. *dat wij malkanderen haest sullen zien met gesontheit
   that we each other soon will see with health

A combination of Christian-ritual elements with contact formulae often occurs when a condition for contact is formulated. In (24e), the contact formula under discussion is used, combined with the health theme through the phrase *met gesontheijt* “with health”,

...
as in (24d). A condition for contact is added to this formula, the essence of this condition being God’s benevolence.

(24) e.  
\[\text{dat wij met gesontheijt malkanderen weer mogen sijen als de goede}\]
\[\text{that we with health each other again may see if the good}\]
\[\text{godt belijeft en ons salijg ijs}\]
\[\text{God+DAT pleases and us blessed is}\]
\[\text{‘that we will see each other again in health, if the good God will like this}\]
\[\text{and will be merciful to us’}\]

This Christian-ritual condition appears only a few times in the letters used for the present chapter. The next section deals with Christian-ritual formulae in general.

8. Christian-ritual formulae

Christian-ritual formulae are often used in connection with intersubjective formulae, where they mainly fulfil a strengthening function (see Sections 4, 5 and 7). Consider one of the subordinate health formulae (see Table 11 in Section 7.1.3) in example (25a). This formula appears in a passage entirely devoted to the health theme, hence we consider it a (subordinate) health formula. At the same time, an appeal is made to God, bringing in the Christian-ritual function.

(25) a.  
\[\text{dat weet Godt Die een kinder daer herten is}\]
\[\text{that knows God who a knower the+GEN hearts is}\]
\[\text{‘that knows God who knows all the hearts’}\]

Greeting formulae and Christian-ritual formulae are usually rather easily separated, mainly because most Christian-ritual formulae fulfil a similar function (see below) to greeting formulae, and are therefore connected to greeting formulae by a conjunction such as \textit{ende} “and”. Such a conjunction enables a clear separation of the different formulae. Nevertheless, what we may call fusion formulae do sometimes occur. Consider example (25b), where the most frequent Christian-ritual function (see below) is interrupted by the greeting \textit{naer cordeale Groetenisse} “after hearty greeting”.

(25) b.  
\[\text{sijt hier mede naer cordeale Groetenisse Godt bevolen}\]
\[\text{be here with after heartily greeting God commended}\]
\[\text{‘with this, be commended into the hand of God, after my heartily greetings’}\]

Similarly, Christian-ritual formulae are often used in combination with contact formulae. In (25c), we present a not uncommon passage from the formulaic ending part of a letter, where intersubjective formulae are combined with Christian-ritual formulae. The passage begins with a greeting (\textit{sijt van harten gegroet} “be heartily greeted”), followed by a Christian-ritual formula (\textit{godt in genaede bevolen} “[be] commended into
the hands of God”). Attached to the antecedent godt “God”, two coordinated relative clauses follow (25c bold). In the first one, the health theme prevails. The second relative clause is a contact formula, expressing the hope that the sender and the addressee will see each other again. In this formula, the health theme is incorporated through the expression met gesontheit “in health” (see Section 7.3). Moreover, since the subclause’s subject die “who” refers to godt “God”, the Christian-ritual theme is also incorporated.

(25) c. sijt van harten gegroet en godt in genaede bevoelen die ick
be from heart greeted and God in mercy commended who I
hoop vl gesontheit en be houwe reis sal geeue en ons
hope you health and safe journey will give and us
met gesontheit bij malkander sal laete koomen
with health to each other will let come
‘be heartily greeted and recommended in the mercy of God who, I hope,
will give you health and a safe journey and who will let us meet in
health.’

While Christian-ritual formulae frequently occur in combination with intersubjective formulae, examples of single Christian-ritual formulae are also found. To gain an impression of the proportion of combined and separate Christian-ritual formulae, we quantified the data from the seventeenth-century subcorpus. In the 81 letters, we find 149 formulaic expressions with overt or covert references to God or with ritualistic language. Of these 149 Christian-ritual formulae, no fewer than 85 tokens, amounting to 57%, occur in combination with intersubjective formulae as in examples (25a–c). The remaining 64 Christian-ritual formulae are for the greater part (51 tokens) variants of one specific formula, viz. the commendation formula.

The commendation formula is commonly found in letters’ endings, its main function being to commend the addressee into the hands of God, just before saying goodbye and closing. In (25d–m), we give representative examples of the commendation formula. In its fullest form, it contains the subjunctive sijt “be” (25d, e, f), the noun godt “God” (25d, g, i) or de heere “the Lord” (25e, h), modified by an adjective such as goede “good” (25d) or aldengooghste “very highest” (25i), the adverbial phrase in ghenade “in grace” (25d, g), and the participle beuolen “commended” (25d, e, g, h). Variation is rampant, however. Reduced forms are very common, with deletion of the subjunctive (25g, h), of the adverbial phrase (25e, f, h) or of the noun, in which case the adjective may be nominalised (25f). The subjunctive may be replaced by indicative forms (25i, j), and the clause can also appear as a subclause (25k). There is phrasal variation: the addressee may be commended to God or the Lord (25d, e, g, h), but also to His protection (25f, i, j, l, m) or mercy (25k). Lexical variation is also found: protection mostly appears as bescherminge (25l), but Romanised forms such as protx, that is, protexie or protectie also appear (25j), as well as bewaringe “lit. keeping” (25m). Finally, the verb bevelen “command” is sometimes replaced by venssen “wish” (25m).
(25)  
d. *sijt defoede* [sic] *godt in ghenade beuolen*
   be the-good God in grace commended

e. *zijt de heere beuolen*
   be the Lord commended

f. *zijt in bescherme des alder hoosten*
   be in protection of-the very highest

g. *godt in genade bevoelen*
   God in grace commended

h. *den heer bevoelen*
   the Lord commended

i. *ick beveel mijn lief en ul altesamen in de bescher mijnge*
   I commend my dear and you all-together in the protection
   des aldenhooghste godt
   of-the very highest God

j. *bevijele ve alle onder protx des aller hoochsten*
   commend you all under protection of-the very highest

k. *dat ul beveele in de ghenade des alder hooghtsen*
   that you commend in the grace of-the very highest

l. *in wiens genadige bescherminge wij u-e altsamen bevelen*
   in whose merciful protection we you all together commend

m. *venssen ve mijn bemijnde in godts bewarenge*
   wish you my dear in Gods protection

With 43 tokens in the 100 eighteenth-century letters selected for the present chapter, the commendation formula remained in frequent use over the course of time.

9. Conclusions

When switching from spoken to written language, letter writers could rely on a fair number of fixed formulae, epistolary formulae which provided them with clear patterns. Text-constitutive formulae offered the means to compose a letter along established lines, to create a text which would immediately be identified as a letter, and to realise a smooth change from one part of the letter to another. Furthermore, intersubjective and Christian-ritual formulae gave letter writers something to hold on to when verbalising experiences. As the sample letter in Section 2 showed, knowledge of the appropriate formulae enabled one to write almost a whole letter. This must have been of extreme importance for people who were not oriented to the practice of writing in their daily life, and not immersed in the written culture. Unsurprisingly, most formulae were quite frequent in the subcorpora used for the present chapter. This bears witness to the
fact that formulaic language was widely used in general, which is also clear from the occurrence of preferred variants and reduced variants. Preferred variants within a variety of similar phrasings give the impression of a converging practice across the language community. Reduced variants typically result from high frequency of use. Finally, the ubiquity of the formulae under discussion in Early and Late Modern Europe became clear from the many resemblances to English and German private letters.

The frequency of formulae in our letter corpora and the ubiquity of these formulae in other contemporary European languages raises the question of how the letter writers learnt to use these formulae. Did they adopt them from the wide range of letter manuals that were frequently reprinted, or did they acquire them by some other means, e.g. in the practice of letter writing within the family? This key question and the possible influence of letter manuals will be addressed in Chapter 5 when we reflect on various aspects of the practice of letter writing. In the present chapter, we have distinguished the text-constitutive, the intersubjective and the Christian-ritual functions of epistolary formulae. In Chapter 4, we will add another function, viz. reduction of the processing effort, and also argue that formulaic language was a ‘safe haven’ for less-experienced writers.
1. Introduction

In the previous chapter, we discussed a large number of fixed formulae in our letters from the seventeenth and eighteenth centuries. These formulae, which fulfill text-constitutive, intersubjective and Christian/ritual functions, show a strong tendency to occur in specific places in the letters, and we established a link between the function of the formulae and text composition. We also focused on similarities among the letters, which appeared to be eye-catching. After having dealt with similarities in Chapter 3, we will concentrate on differences in the present chapter, and discuss both synchronic variation in the use of formulae and diachronic changes from the seventeenth to the eighteenth century, taking into account the variables of gender, class, and time.

An explanation for the variation and change found will be given from different perspectives. First and foremost, we consider the variation and change in formulaic language in connection with literacy and writing experience. For less-experienced writers, epistolary formulae may offer a conventionalised ‘safe option’. Like other lexicalised multi-word units, they can be retrieved as a whole from memory, which would make writing easier and thus speed up the writing process. If the use of formulae indeed reduces the writing effort, it is expected that writers suffering formulating problems will prefer fixed formulae which can compensate for a lack of writing experience and (compositional) creativity. This function of formulae as a ‘safe option’ has not yet been discussed in the previous chapter. When discussing writing experience, we have to realise that literacy and writing experience were socially distributed, and that the literacy rate increased from the seventeenth to the eighteenth century. We expect, therefore, that both synchronic and diachronic distributional effects in the use of formulae could at least in part be accounted for by the social distribution of literacy and writing experience. In other words, we expect to find more formulae in letters written by less-experienced writers.

There is, however, another explanation that we have to take into account here. Differences between social groups originating from relatively less and more writing experience may become markers of group languages. If, for instance, less-experienced writers use more epistolary formulae, these formulae may become characteristic of their language and the language of the social groups with which they identify (cf. Laitinen & Nordlund 2012). This may apply even to experienced writers from these social groups, and even when writing experience increases over time. The use of formulae may thus develop into a group-related common practice, independent of the
level of writing experience. Considering this possibility, we will nevertheless argue that
the variational patterns found favour an explanation in terms of writing experience
rather than identity roles or group practices.

In summary, we will discuss the following three topics:

– How does the use of epistolary formulae vary across social class, gender, and
time?
– Can we explain the variation and change found in the use of epistolary formulae
with an appeal to writing experience?
– Can we explain the variation and change found in the use of epistolary formulae
with an appeal to identity roles or group practices?

In Section 2 of the present chapter, we will discuss the link between the functions of
formulaic language, writing experience, sociolinguistic variation and diachronic
change, which will lead us to several hypotheses. In Section 3, we will present the re-
sults from a series of case studies into specific formulae. The corpus used for these case
studies is described in Section 3.1. In the conclusion of Section 3.8, we discuss why we
consider writing experience to be the overarching factor in explaining the variational
patterns found in Section 3. In Section 4, we will continue discussing writing experi-
ence and will also consider alternative explanations. Section 5 contains the conclu-
sions and further discussion of a number of questions that arise.

2. Formulaic language and writing experience

2.1 Previous research

Formulaic language has received much scholarly attention from various perspectives
and frameworks. Apart from research on epistolary formulae, formulaic language has
been studied from a predominantly psycholinguistic point of view (Wray 2002), and
from a sociolinguistic and ethnographic point of view (Kuiper and associates, see
Kuiper 2009), as well as from conversation-analytic (Tannen 1987; Norrick 2000),
historical-linguistic (Corrigan et al. 2009) and pragmatic perspectives (Coulmas 1979,
1981a). Moreover, similar topics have been addressed in phraseology (Cowie 1998;
Granger & Meunier 2008), and are also at the core of constructional approaches to
grammar (Goldberg 1995, 2006), pattern grammar (Hunston & Francis 2000), and
corpus linguistics (Biber 2009). This is not to say that, for instance, constructions in
the sense of construction grammar and epistolary formulae are one and the same. All
these approaches, however, share an interest in multi-word strings or so-called ex-
tended lexical units (Sinclair 1998, 2008; Poss & Van der Wouden 2005), and regard a
strict separation of grammar and lexicon as problematic.

Questions that have been addressed in earlier research into epistolary formulae in
historical letter-writing include such topics as power and politeness (Nevalainen &
Raumolin-Brunberg 1995; Dossena 2003; Nevala 2004b; Tiisala 2004; Bax 2010), sociolinguistic variation (Tieken-Boon van Ostade 1999; Austin 2004), the relation between formulaic and expressive language (Wood 2009), and the functions of epistolary formulae (Elspaß 2005: 152–196). In a previous study, using a smaller number of letters than for the present chapter, we were able to establish sociolinguistic variation, i.e. both gender and social patterns, in the use of epistolary formulae (Rutten & van der Wal 2012). A gender analysis of the seventeenth-century material revealed that, on average, women used far more formulae than men. For the eighteenth-century letters, a clear social pattern was shown in that towards the end of the 1700s, the formulae under discussion were mainly used by members of the lower (middle) classes, and hardly at all in the upper (middle) classes. In Rutten & van der Wal (2012), we related these results to the social distribution of literacy, schooling and writing experience. Furthermore, the gender and social distributions were in line with the findings of Austin (2004), who studied the decline of epistolary formulae in English letters from the eighteenth and nineteenth centuries, and who concluded that “[t]he two main groups that continue to use the formulas, even into the nineteenth century, are seamen, mostly of the lower ranks, and women”, in other words, social and gender factors explained the distribution of formulae over the language community.

As well as dealing with sociolinguistic variation in the present chapter, we also continue our focus on the functions of formulaic language. In Chapter 3, we showed which epistolary formulae fulfilled the text-constitutive, inter-subjective and Christian-ritual functions in our letters. We also indicated (Chapter 3, Section 9) that we might need to distinguish a fourth function, viz. reduction of the processing effort. This fourth function will be at the core of the present chapter.

2.2 Processing effort and writing experience

Wray (2002: 101), who offers a comprehensive overview of the functions of formulaic language, collapses the many possible functions of this type of language into three main functions: “the reduction of the speaker’s processing effort, the manipulation of the hearer (including the hearer’s perception of the speaker’s identity), and the marking of discourse structure”, in short: “processing, interaction and discourse marking.” The inter-subjective, Christian-ritual and text-constitutive functions discussed in Chapter 3 correspond to interaction and discourse marking. These interactional and discourse marking functions in present-day spoken language can thus be applied to historical written language: they were straightforwardly found in Dutch letters from the seventeenth and eighteenth centuries. Wray’s third function,

1. For the previous study, our corpus consisted of 81 autographs from the seventeenth century, and 100 from the eighteenth century.
2. These three main functions are then “further reduced to one overriding priority, the speaker’s promotion of self” (Wray 2002: 101).
Letters as Loot

processing, is a psycholinguistic notion, pointing to the speaker’s ability “to produce language in a more efficient, less energy-draining way” (Tannen 1987: 581). Using prefabricated strings, the speaker buys time for processing, and creates a shorter processing route (Wray 2002: 97). In short, retrieving formulae from memory is relatively easy, in particular in standard communicative situations (Coulmas 1981b). The only effort a language user needs to make is to memorise the formula. When this is done, the formula can time and again be retrieved from memory as a whole unit, without the need for the language user to compose a novel formulation. At first sight, the reduction of the processing function, characteristic of on-line speech production, is not a self-evident phenomenon in a historical corpus of written texts, but closer examination will show otherwise.

How could an advantage in on-line language production be applied to written language? Although the difference between speaking and writing is not absolute (Biber 1988; Koch & Oesterreicher 1985), which is in fact presupposed in much historical sociolinguistic work (cf. Elspaß 2005), there is an absolute difference with regard to time, and therefore with regard to processing. In conversation, production and reception take place virtually at the same time, and the use of formulae is, therefore, efficient, for both the speaker and the hearer. Written communication, however, does not suffer from time pressure, and is conversely characterised by “the property of being worked over” (Chafe 1994: 43). Writers may adjust, rephrase, edit or rewrite anything they have written, thus buying, in a sense, additional processing time. As a consequence, the advantage of formulaic language disappears in as far as it is connected to efficiency and reduction of the processing effort.

Nonetheless, it might be argued that reduction of the processing effort is also important in the case of written language, adopting Elspaß’s (2005: 180–181) idea that formulaic language may function as Formulierungshilfe “formulating help”. Recourse to formulae would not only serve text-constitutive and inter-subjective or Christian-ritual functions, but would also provide a safe option for writers suffering problems of formulation. People who were not used to writing regularly, and who were suddenly forced to write because family members or they themselves were at a geographical distance, in the country or abroad, may have been faced with a difficult task. In such cases, the use of lexicalised formulae, retrieved as whole units from memory, would have made writing easier and speeded up the writing process. As a result, we would expect less-experienced writers to need and to use more formulae.

The importance of writing experience in historical sociolinguistics, often operationalised with reference to literacy, schooling and the writers’ daily routine, has been put forward by Vandenbussche (1999) and Elspaß (2005: 45–46) among others. The writers’ daily routine refers to their contact with specific varieties in their everyday life and/or profession, as a crucial distinction can be made between writers who are mainly handarbeiterorientiert “oriented to manual labour”, and others who are partly or largely schriftarbeiterorientiert “oriented to writing” (Vandenbussche 2002: 39). Specifically concerning formulaic language in historical letters, Elspaß (2005: 192) claimed that
formulaic language was predominantly used by inexperienced writers when solving communicative problems in the written code. Instead of pondering for a long time, the writer could resort to fixed formulae that provided conventionalised and generally accepted ways of verbalising information and experiences. Kuiper & Haggo (1984: 224), discussing the formulaic language of present-day livestock auctioneers, compared the process of becoming a fluent auctioneer to the oral poet’s transition from a neophyte to a young singer, and finally to a mature singer. The acquisition and production of auctioneers’ formulae depends on experience, and the process runs from learning the formulae from an experienced practitioner, through the use of “established formulae in established ways” (1984: 224) to creativity. The second stage is characterised by the use of conventionalised formulaic language, the third by creative variation on established patterns. Similarly, we would expect the most experienced letter writers, who had more contact with written varieties thanks to schooling and/or daily activities, to be the most creative writers, using the fewest formulae.

One of the goals of the present chapter is to test whether such a link between writing experience and formulaic language can indeed be established. The question, then, is who were the most literate and most experienced writers in the seventeenth and eighteenth centuries and who were the least literate and the least experienced?

### 2.3 Writing experience and social variables

With regard to the second half of the seventeenth century, it is estimated that two-thirds of the male population and one-third of the female population in the Northern Netherlands were able to write (Frijhoff & Spies 1999: 237). It should be noted that literacy rates are usually estimated by counting signatures in marriage registers, although being able to sign, that is, to write one’s own name, does not necessarily imply full writing skills. Actual rates of full literacy may therefore have been lower. Furthermore, we would like to stress that literacy is a gradual notion, not a discrete one and that there are various stages between being fully illiterate and having perfect writing skills (cf. Fairman 2007). The estimated literacy rates show gender variation, originating from gender-specific schooling opportunities and/or gender-specific daily activities involving or not involving writing experience. Housekeeping, for instance, a frequent female activity does not necessarily require writing. Around 1800, the average literacy

---

3. In this case, it should also be noted that reading and writing were commonly taught consecutively, so that the number of skilful readers may have been higher than the number of writers. Moreover, writing was more expensive than reading because of the costs of materials. De Booy (1977: 55), discussing educational practices in the province of Utrecht in the seventeenth and eighteenth centuries, notes that many parents ended their children’s school careers when writing came in sight, for financial reasons, but also because children of eight or nine years old were potential workers.
scores for both men and women had increased to 80% and 60% respectively, while the gender differences remained (Kloek & Mijnhardt 2001: 81).

Literacy appears to have been not only gendered, but also socially stratified in that one-third of the two lower ranks and two-thirds of the other ranks were literate in the later seventeenth century (Frijhoff & Spies 1999: 238). We distinguish between four social ranks (lower, lower-middle, upper-middle, upper; see Chapter 1, Table 1) following the stratification commonly made by historians (Frijhoff & Spies 1999: 188–190; Bruijn 2008: 16). These social strata are primarily defined in terms of profession and/or occupation. For a further discussion of this social ranking we refer to Chapter 1. It should be noted that to assign specific letter writers to social ranks or classes, we used a variety of criteria, the most important of which is the writer’s profession or, in the case of women, the husband’s profession.

If we approach writing experience in terms of literacy and schooling or profession, this leads us to three hypotheses, the first suggesting a gender distribution, the second a social and the third a diachronic distribution. The gender distribution of seventeenth- and eighteenth-century literacy leads us to expect gender variation in the use of formulae. As our social stratification is largely founded upon the profession of the writers, and as schooling opportunities were socially stratified (Frijhoff & Spies 1999: 238), we assume that our social stratification runs parallel to the level of schooling, and to everyday involvement in written varieties. Our second hypothesis would therefore be that the higher social ranks use formulaic language to a lesser extent than the lower social strata. The increasing literacy rates from the seventeenth to the eighteenth centuries point to growing schooling opportunities, and suggest a wider spread of (knowledge of) written varieties of Dutch. Diachronically, we therefore expect a decrease in the use of formulaic language.

2.4 Summing up – hypotheses

In sum, building on Wray (2002), Austin (1973a, 1973b, 2004), Elspaß (2005) and Rutten & van der Wal (2012), we formulate the following hypotheses.

– We expect to find more formulaic language in letters by women than in letters by men.

4. We are aware of the fact that class is a nineteenth-century category, although there were, of course, social divisions or strata before the nineteenth century. We use the common sociolinguistic term class to refer to these pre-nineteenth-century social divisions.

5. Frijhoff & Spies (1999: 237) found hardly any regional differences of average literacy rates, although there was a link between urbanisation and literacy (Frijhoff & Spies 1999: 238), as average literacy rates were higher in cities than in the countryside. A possible discrepancy between city and countryside cannot be examined in the present letter material, which consists of letters mainly from and to towns and cities in Holland (Amsterdam, Enkhuizen, Hoorn, Rotterdam) and Zeeland (Vlissingen and Middelburg).
– We expect to find more formulaic language in letters from the lower classes than in letters from the upper classes.
– We expect to find more formulaic language in letters from the seventeenth century than in letters from the eighteenth century.

In the next section, we report on a series of case studies to see if these expectations are fulfilled.

3. Case studies

This section will examine the patterns of variation and change in a large number of epistolary formulae. In the case studies presented in Sections 3.2 to 3.7, we will primarily be concerned with the link between the use of epistolary formulae and the external variables of social class, gender and time. We selected intersubjective as well as Christian-ritual and text-constitutive formulae, all of which were highly frequent (see Chapter 3 and see below) in order to achieve results as reliable as possible. Although we will occasionally diverge from our quantitative approach and discuss qualitative data, the main goal of this chapter is to quantitatively establish variational patterns in the corpus as a whole, by the use of independent sociolinguistic variables. Thus, our main goal is to establish possible links between the use of formulae and the language found in particular social groups. Whether these links are due to writing experience or not, will be discussed in Sections 3.8 and 4.

Firstly, we will discuss the greeting formula *een vriendelijke groetenisse zij geschreven aan* “a friendly greeting be written to” (Section 3.2). Then, we will focus on a number of health formulae and subordinate health formulae (3.3–3.5). Finally, we will discuss two text-structural formulae: *ik laat u weten als dat* “I let you know that”, which indicates a new topic or a new part of the letter (3.6) and *voorts/verder* “further”, which either continues a topic or indicates a new one (3.7). We deliberately did not take into consideration addresses, dates and other elements identifying the document as a letter, as these offer the writer less choice (see Chapter 3, Section 4). Obviously, pragmatic or sociolinguistic factors may influence the choice of address terms (cf. Bax 2010; Tiisala 2004; Nevala 2004b; Tieken-Boon van Ostade 1999; Nevalainen & Raumolin-Brunberg 1995; Dossena 2003), but here we are less interested in politeness-theoretical and other socio-pragmatic aspects of the choice of variants than in the choice of not using any variant at all. When the writer chooses not to use any variant, in this case a formula, at all, variation and change also become manifest. Such variation appears to be more interesting from the perspective of processing effort and writing experience, and is also of interest in view of our concern with the writing process. Before embarking on the case studies, we will introduce the corpus (3.1).
3.1 The subcorpora

For our analyses we used the corpus presented in Chapter 1, Section 6, more in particular the seventeenth-century autograph subcorpus and the total eighteenth-century subcorpus which consists of only autographs (Simons 2013: 52–55; 87–93). The basic numbers of the subcorpora (Chapter 1, Table 2) are repeated here in Table 1. The data for the seventeenth-century autograph corpus, which is part of the seventeenth-century total subcorpus, are presented separately between brackets.

The composition of the subcorpora in terms of **gender** and **social class** is given in Tables 2 and 3, divided horizontally into four social classes, and vertically into male and female writers. The social class index labels are *lower class* (LC), *lower middle class* (LMC), *upper middle class* (UMC), and *upper class* (UC); see Chapter 1 for an explanation of these social classes. For each social class and for each gender, the number of letters (N Le) and the number of words (N Wo) are given. M refers to male writers, F to female writers. Note that we were able to establish the social class of 219 from the total of 260 seventeenth-century autographs.

**Table 1.** Basic numbers of the subcorpora used.

<table>
<thead>
<tr>
<th>Period</th>
<th>Subcorpus</th>
<th>N letters</th>
<th>N words</th>
<th>N writers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1660s-1670s</td>
<td>Total</td>
<td>549</td>
<td>228,000</td>
<td>424</td>
</tr>
<tr>
<td>(1660s-1670s)</td>
<td>Autograph</td>
<td>260</td>
<td>118,000</td>
<td>202</td>
</tr>
<tr>
<td>1770s-1780s</td>
<td>Total</td>
<td>384</td>
<td>196,500</td>
<td>292</td>
</tr>
</tbody>
</table>

**Table 2.** The seventeenth-century autograph subcorpus by social class and gender.

<table>
<thead>
<tr>
<th>17th century</th>
<th>LC</th>
<th>LMC</th>
<th>UMC</th>
<th>UC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Le</td>
<td>N Wo</td>
<td>N Le</td>
<td>N Wo</td>
<td>N Le</td>
<td>N Wo</td>
</tr>
<tr>
<td>M</td>
<td>5</td>
<td>3,000</td>
<td>21</td>
<td>11,000</td>
<td>121</td>
</tr>
<tr>
<td>F</td>
<td>5</td>
<td>2,500</td>
<td>20</td>
<td>11,000</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>5,500</td>
<td>41</td>
<td>22,000</td>
<td>145</td>
</tr>
</tbody>
</table>

**Table 3.** The eighteenth-century subcorpus by social class and gender.

<table>
<thead>
<tr>
<th>18th century</th>
<th>LC</th>
<th>LMC</th>
<th>UMC</th>
<th>UC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Le</td>
<td>N Wo</td>
<td>N Le</td>
<td>N Wo</td>
<td>N Le</td>
<td>N Wo</td>
</tr>
<tr>
<td>M</td>
<td>24</td>
<td>8,000</td>
<td>83</td>
<td>28,000</td>
<td>93</td>
</tr>
<tr>
<td>F</td>
<td>2</td>
<td>1,000</td>
<td>14</td>
<td>6,000</td>
<td>38</td>
</tr>
<tr>
<td>Total</td>
<td>26</td>
<td>9,000</td>
<td>97</td>
<td>34,000</td>
<td>131</td>
</tr>
</tbody>
</table>
Table 4. The subcorpora used for the present case studies.

<table>
<thead>
<tr>
<th>Period</th>
<th>N letters</th>
<th>N words</th>
<th>Av N words/letter</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th century</td>
<td>219</td>
<td>102,000</td>
<td>466</td>
</tr>
<tr>
<td>18th century</td>
<td>384</td>
<td>196,500</td>
<td>512</td>
</tr>
</tbody>
</table>

It is not possible to provide consistent data for all the cells in the tables. For the eighteenth century in particular, we regret the low number for women in the LC, where the number of two letters, amounting to 1,000 words, forces us to be wary of drawing too firm conclusions.

Since we are not dealing with traditional sociolinguistic variables, we have to rely on counting frequencies per 10,000 words in the case studies in Sections 3.2 to 3.7. This enables us to compare the relative frequencies of the different formulae. Some formulae typically appear only once per letter, such as the greeting formula *een vriendelijke groetenisse zij geschreven aan* “a friendly greeting be written to” (see Section 3.2). In such cases, we will still count per 10,000 words for reasons of comparability, but we will also investigate the numbers of letters with this formula. Table 4 presents the subcorpora and their average number of words per letter, showing that there is only a slightly higher average number of words for the eighteenth-century than for the seventeenth-century letters.

3.2 Greeting

As discussed in Section 6.1.3 of Chapter 3, many seventeenth-century private letters begin with the formula *een vriendelijke groetenisse zij geschreven aan* “a friendly greeting be written to”. A full example is given in (1a). Since the formula exhibits some considerable variation, we will first explain what counts as a token of the formula, and provide ample examples (1a–1i). Fixed elements of the formula are the article *een* “a”, the noun *groetenisse* “greeting” and the preposition *aan* “to”. An alternative to the most common adjective *vriendelijke* “friendly” is *hertelijke* “hearty, cordial” (1b). We also found one instance of *hartgrondige* “whole-hearted” (1c). In very rare cases, the adjective is omitted (1d) and in some cases it is preceded by an intensifying adverb such as *seer* “very” (1e). The verbal forms *zij ‘be’ or geschreven ‘written’ may be left out (1c, 1g), and sometimes both are omitted (1h). Morphological or phonological variation is mainly found in the subjunctive *zij* or *sij ‘be’, which is also found as *zijn* or *sijn* (1g), in the noun *groetenisse*, where the final syllable may be deleted (1i), and in the preposition *aan* or *aen “to”, which may be shortened to *an* (1h). All three variants are familiar features of north-western dialects, spoken in the Holland and Zeeland areas.

(1) a. *een vrendelycke groetenyese sy geschreven aan myen lyve en ser*  
a friendly greeting be written to my dear and very
bemyen de maen pyeter eversen  
beloved husband Pyter Eversen 
b. een hertelike groetenisse sij geschreven aen min lieven en wel  
a hearty greeting be written to my dear and very  
beminde man met name pieter carelsen  
beloved husband with name Pieter Carelsen 
c. Een hartgrondige groetenisse geschreven aen mijn seer Beminde  
a whole-hearted greeting written to my very beloved  
broeder Jan harmenss  
brother Jan Harmenss 
d. een groetenisse sij gecrewen aen mijn beminde maen derck  
a greeting be written to my beloved husband Derck  
ahbramsoen  
Ahbramsoen 
e. een seer vriindelijck groitenisse sijn geschreven aen mijn seer  
a very friendly greeting be written to my very  
beminde aldelliste man ijan heije  
beloved dearest husband Yan Heije 
f. een vriendeliche groetnisse geschreuen an min seer beminde broeder  
a friendly greeting written to my very beloved brother 
g. Een vriindelick grotenis sin aen vl min seer beminde vrouwe  
a friendly greeting be to you my very beloved wife  
gretien Jansen 
Gretien Jansen 
h. een vriendelijke grotenjse an mijn seer bemende man sijewert  
a friendly greeting to my very beloved husband Sijewert  
leenders  
Leenders 
i. een vriendelijck grotenes aen ul mijn beminde soon roelof barendts  
a friendly greeting to you my beloved son Roelof Barendts 

Extracting all relevant examples from the 219 letters of the seventeenth-century autograph subcorpus resulted in 42 instances of this formula. Figure 1 shows the distribution by social class. What we see is a strong decrease towards the right-hand side. In other words: the higher the social class index, the lower the frequency of the formula. 

Examining the distribution by gender we also find a clear pattern of variation which is presented in Figure 2. Women use the formula een vriendelijke groetenisse... more frequently than men do.
Chapter 4. Variation and change in formulaic language

Figure 1. Frequency per 10,000 words and by social class of *een vriendelijke groetenisse...* “a friendly greeting...” in the seventeenth century.

Cross-tabulating class and gender shows that the gender pattern holds across social class (Figure 3). Low numbers force us to combine the LC+LMC (24 tokens) and the UMC+UC (18 tokens). The black columns represent the frequency in letters from men, the grey columns the letters from women. The general gender difference in Figure 2 appears to hold for the two lower and two upper classes separately: women still use the formula more frequently than men. Furthermore, the combined social division as in Figure 3 clearly shows the same social class effect as represented in Figure 1: the higher the social class index, the lower the frequency of the formula. The social ranks display, firstly, a decrease, and secondly, a consistently lower frequency for men than for women.

Figure 2. Frequency per 10,000 words and by gender of *een vriendelijke groetenisse...* “a friendly greeting...” in the seventeenth century.
For this formula, which typically occurs only once per letter, viz. as the opening line, we also counted the occurrence per letter (cf. Table 5). We find the same pattern of social and gender variation as before. About half of the LC and LMC letters comprise the formula and 43% of the letters by women, while much lower proportions of 12% and 4% respectively are found in the UMC and the UC letters and only 11% in letters written by men. Figure 4 presents the results of cross-tabulating class and gender and combining LC and LMC (24 tokens) against UMC and UC (18 tokens). Figure 4 confirms the earlier pattern (Figure 3): gender differences hold within the social classes.

Turning to the last external variable, we witness a very strong diachronic effect on the use of the formula *een vriendelijke groetenisse zij geschreven aan* “a friendly greeting be written to”: not a single variant of the formula was found in the eighteenth-century subcorpus. In the opening of letters, no greetings are communicated (Chapter 3, Section 6.1.3). We often find greetings further on in the letters (Chapter 3, Section 7.2), but the formula studied here seems to have disappeared from use. This could be
Chapter 4. Variation and change in formulaic language

Figure 4. Proportion of letters with the formula *een vriendelijke groetenisse...* “a friendly greeting...” in the seventeenth century, by social class and by gender.

concluded to be a classic example of a change from above, with U(M)C men leading the change. We will now see whether similar developments occurred in the case of other formulae.

3.3 Communicating health

As we saw in Chapter 3, the statement *dat ik en X nog kloek en gezond zijn* “that I and X are still strong and healthy” is one of the most frequent formulae in seventeenth-century letters. The statement is usually made in the first part of a letter, somewhere in the initial lines. The subject is most often first person singular or plural (2a, 2b). There are also cases with a second person subject, where the writer expresses the hope that the addressee is still healthy, or refers to an earlier health statement by the addressee (2c). Third person subjects also occur, especially when the formula is not used in the first part of a letter, but further on, where the health of third parties is communicated (2d).

(2) a. *mijn bemijnde broeder Jan petijt ick laet u.l. weten als dat ick noch klouck ende gesont ben* my beloved brother Jan Petijt I let you know as that I still strong and healthy am

b. *ick laet eu weten als dat wi al te samen noch klock en gesondt zijn* I let you know as that we all together still strong and healthy are
c. daer uijt uer staen daet u l noch cloec en gesont waes
there from understood that you still strong and healthy were
'[and I have] understood [from your letter] that you were still strong and
healthy'

d. Wat aengaet van vader is noch clouck ende gesonth
what concerns of father is still strong and healthy
‘concerning father, [he] is still strong and healthy’

Since the subject may vary, there is also variation in the verbal forms following kloek
en gezond “strong and healthy”, depending on person, number, time and dialect. We
consider kloek en gezond as the key words of the formula. This is a fixed combination
which occurs many times, and in many spellings, such as kloeck, klock, cloec, clouck,
gesont, ghesond, etc.

We extracted all relevant forms from the 219 seventeenth-century autographs,
which led to 52 tokens. Only the combination kloek en gezond counts as a token. We
thus excluded an example with clock ende leersaem ‘strong and inquisitive’, a token with
the adjectives nominalized (wy wensen u ende Altesamen cloeckheyt ende gesont heit
‘we wish you all strength and health’), and 21 tokens with just kloek. As this formula
may appear many times in one letter, and not typically only once as the previous for-
mula (Section 3.2), we did not count the number of formulae per letter, but only per
10,000 words. Figure 5 plots the frequency per 10,000 words by social class.

Whereas the lower class scores approximately 12 occurrences per 10,000 words,
the frequency drops to approximately 6 in the lower middle class and 4 in the upper
middle and upper classes. In other words, we witness a clear decrease of the formula
when the social class index rises.

![Figure 5](image)

**Figure 5.** Frequency per 10,000 words and by social class of kloek en gezond “strong and
healthy” in the seventeenth century.
Gender variation is attested here as well, as is shown in Figure 6. Women use the *kloek en gezond* formula almost twice as frequently compared to men. Cross-tabulating class and gender and again combining the LC and LMC (21 tokens) and the UMC and UC (31 tokens) results into Figure 7. Gender variation is clearly attested in the L(M)C, but no gender difference is found in the U(M)C.

Apart from gender and social distributions, we also find a strong diachronic effect. Not a single instance of *kloek en gezond* was found in the eighteenth-century letters. Similarly, we did not find any tokens of *een vriendelijke groetenisse zij geschreven aan*...
“a friendly greeting be written to” (Section 3.2). This is, however, not the only observation to be made on diachronic change. Instead of kloek en gezond, we find another equally tautological and fixed expression fris en gezond “lit. fresh and healthy, in good health” in the eighteenth-century subcorpus, which formula is not attested in the seventeenth-century autograph subcorpus. This expression mainly occurs in the first part of letters (2e, f), as kloek en gezond did in the seventeenth century, but sometimes also further down, after greetings and the contact formulae, as in (2g):

(2) e. ik laat uwe weeten als dat ik nog fris en gesont ben
   I let you know as that I still fresh and healthy am
   ‘I let you know that I am still in good health’

f. ik laat u alle weten als dat ik godt zy lof nog fris en gesondt ben
   I let you all know as that I God be praise still fresh and healthy am
   ‘I let you all know that I, praise God, am still in good health’

g. wat myn en myn kinderen aangaat wy zyn door de goedheyd Gods alle nog fris en gesond
   what me and my children concerns we are by the goodness God’s all still fresh and healthy
   ‘as far as me and my children are concerned, we are all still in good health by the grace of God’

We extracted all the instances of this expression from the eighteenth-century subcorpus, which resulted in 40 tokens of fris en gezond occurring in the 384 letters of the subcorpus. Before turning to the diachronic difference, we first show the distribution of the formula across social class in Figure 8.

Figure 8. Frequency per 10,000 words and by social class of fris en gezond “fresh and healthy” in the eighteenth century.
The social class pattern found earlier also appears in the eighteenth-century material. The columns sharply decrease towards the right-hand side of Figure 8, which means that the formula occurs less frequently in the lower middle and in the upper middle and upper classes. A clear gender difference, however, does not appear: men score 2.3 tokens per 10,000 words, women 1.3 tokens per 10,000 words. This gender levelling is confirmed when cross-tabulating (and combining, as before) the frequencies of L(M)C (27 tokens) and U(M)C (13 tokens), as the fairly similar gender scores in Figure 9 show.

![Figure 9](image.png)

**Figure 9.** Frequency per 10,000 words, by social class and by gender of *fris en gezond* “fresh and healthy” in the eighteenth century.

Turning to the diachronic differences, we compare the total scores of the two health formulae discussed above, viz. the seventeenth-century formulae *kloek en gezond*, and its eighteenth-century successor *fris en gezond*. Figure 10 presents the results, and is based on the total scores represented in Figures 5 and 8 above. The formula changes, *kloek* disappearing and *fris* appearing, but diachronically the decrease towards the right-hand side is present in both cases. Apart from the social class effect, the diachronic difference is also clear: the grey columns representing the eighteenth-century usage of *fris en gezond* are consistently lower than the black columns representing the seventeenth-century usage of *kloek en gezond*, with the only exception of the LC. This means that in most social classes the health formulae under discussion are more frequent in the seventeenth century than in the eighteenth century. We conclude that there is indeed a diachronic difference, since the use of health statements decreases over time, and that this development is particularly strong in the middle classes and in the UC, and to a lesser extent in the LC.
In this section, we discuss a so-called subordinate health formula (see Chapter 3, Section 7.1.3) that often appears in the very formulaic first part of the letters. This means that it usually follows health statements or health wishes, expressed by such formulae as discussed in the previous section. After having expressed the wish that the addressee is still in good health, many writers remark that if that were not the case, they would be sorry to hear that (3a).

(3) a. waer het anders het soude mijn van herten leet sijn om
were it different it would me from heart sorrow be to
'thooren
to hear
‘if it were different, I would be very sorry to hear that’

The key word in this formula is the noun leed “sorrow”. This is usually preceded by van herten “from the heart, heartily” (3a), sometimes by seeer “very” (3b), or by both (3c), or by nothing (3d). The noun leed is combined with the subjunctive het ware “it were” (3d), or the modal auxiliary sou[de] “would” plus the infinitive of either zijn “be” (3a–c) or wesen “be” (3e). In almost all cases, the formula continues with om te horen “to hear”, although occasionally this is omitted (3b, e). As a rule, the conditional construction ware het anders “if it were different” precedes the formula (3a, c–e). The condition is sometimes rendered with a different conditional construction (3b).

(3) b. soo anders ware ‘tsoude ons alle seeer leet sijn
if different were it would us all very sorrow be
‘if it were different, we would all be very sorry’
c. *waer het anders het souden mijn seer van harten leet zijn*  
were it different it would me very from heart sorrow be  
*om te hooren*  
to to hear  
‘if it were different with you, we would be very sorry to hear that’

d. *waer het anders het waer mijn Leedt om hooren*  
were it different it were me sorrow to hear  
‘if it were different, I would be sorry to hear’

e. *ware het anders het soude mij van herten leet wesen*  
were it different it would me from heart sorrow be  
‘if it were different, I would be very sorry to hear’

For our study of the external variables, we first extracted all tokens of this formula from the seventeenth-century autograph subcorpus: 55 tokens in the 219 letters. Figure 11 presents the frequency of this formula per 10,000 words, divided by social class. Again, a clear social distribution is shown, steadily decreasing from approximately 9 in the LC, 7,5 in the LMC and 5 in the UMC to approximately 1,5 tokens per 10,000 words in the UC. We also see the familiar gender pattern with a higher frequency for women than for men as becomes clear in Figure 12. Cross-tabulating class and gender, we notice that the gender difference clearly holds for the L(M)C, but is absent from the U(M)C where the scores are almost equal (cf. Figure 13).

Extracting all the instances of the formula from the eighteenth-century subcorpus led to 89 tokens in 384 letters, which were subsequently allocated to the social and gender variable. Figure 14, plotting the frequency per 10,000 words and by social class shows that the social pattern attested for the seventeenth century also holds for the eighteenth century. It should be noted that hardly any instances are attested for the UC. Remarkably, the eighteenth-century frequencies in the LC and LMC are higher than the seventeenth-century scores.

![Figure 11](image-url)

**Figure 11.** Frequency per 10,000 words and by social class of *van herten leed* “very sorry” in the seventeenth century.
Figure 12. Frequency per 10,000 words and by gender of *van herten leed* “very sorry” in the seventeenth century.

Figure 13. Frequency per 10,000 words, by social class and by gender of *van herten leed* “very sorry” in the seventeenth century.

Figure 14. Frequency per 10,000 words and by social class of *van herten leed* “very sorry” in the eighteenth century.
Figure 15. Frequency per 10,000 words and by gender of *van herten leed* “very sorry” in the eighteenth century.

Figure 16. Frequency per 10,000 words and by social class of *van herten leed* “very sorry” in the eighteenth century.

Figure 15 reveals that for the eighteenth century the gender effect commonly found in the use of epistolary formulae, viz. that women use more formulae than men, is reversed. Cross-tabulating social class and gender results into a similar reversed pattern for the L(M)C. For the U(M)C, Figure 16 shows the same levelling as in the seventeenth century (cf. Figure 13).

Comparing the results for the seventeenth and the eighteenth centuries, a remarkable difference is established which was already mentioned above: the rising frequency for the LC and LMC. The patterns are visualised in Figure 17, which combines the scores from Figures 11 and 14.
Letters as Loot

For the UMC and UC the diachronic difference is what we would expect: a lower frequency in the eighteenth century than in the seventeenth century. The black columns, representing seventeenth-century usage decrease towards the right-hand side, and in the eighteenth-century (the grey columns), the formula has almost completely disappeared from UC usage. In the LC and LMC, on the other hand, the eighteenth-century frequency remains quite high, and is even higher than in the seventeenth century. This is, however, not a sign of a general increase in the use of this formula, but only a sign that it remained more in use in the lower classes than in the upper classes.

3.5 Communicating God’s omniscience

In this section, we discuss another subordinate health formula (Chapter 3, Section 7.1.3), notably one which expresses the sender’s regret, if he were to be informed that the addressee were no longer in good health. This regret is expressed with reference to the omniscience of God, where the appeal to God is intended to render the sender’s health wish and regret as genuine as possible. The formula often follows the formula van herten leed “very sorry” discussed in Section 3.4. It is frequently part of a formulaic opening, as exemplified in (4a).

(4) a. *hoope dat het met u mijn lieve man oock soo is waer het* hope that it with you my dear husband also so is were it
   *anders het soude mij van harten* different it would me from heart
   *leet wesen om te hooren dat weet godt almaghtig die een* sorrow be to to hear that knows God almighty who an
**kender der harten is**

**expert the+gen hearts is**

‘[I] hope that you, my dear husband, are also like that [viz. healthy]. If it were different, I would very much regret to hear that. This knows God Almighty, who knows all the hearts’

The formula is introduced by, and consists in its most rudimentary form of *dat weet God* “that knows God”. The name of God can be modified with an adjective such as *goede* “good” (4b), with an addition such as *almachtig* “almighty (4a, 4c). The relative clause almost always mentions God as an expert on the hearts of men, sometimes expressed with a prepositional phrase such as *van al herten* “of all hearts” (4c), sometimes with a genitive (4a). Some variation (4b) may occur, or God is simply presented as omniscient (4d), or another quality of God is added (4e).

(4) b. *dat weet de goeden godt die een kenner is van mijn haerdt*

that knows the good God who an expert is of my heart

*ende van alle menschen haerten*

and of all men hearts

‘that knows the good God who is an expert of my heart and of all men’s hearts’

c. *dat weet god almachtig de is een kender van al herten*

that knows God almighty who is an expert of all hearts

‘that knows God Almighty who is an expert of all hearts’

d. *dat weet godt Dije een kender alles Is*

that knows God who an expert all is

‘that knows God who is an expert of everything’

e. *dat weet godt die onze herte en nieren Cent*

that knows God who our hearts and kidneys knows

‘that knows God who is an expert of the hearts and of the kidneys’

We find 29 tokens of this formula in the seventeenth-century autograph subcorpus. Caution should be taken because of the fairly low numbers, but still Figure 18 gives the familiar distribution pattern of the formula across social class, as before counted per 10,000 words. The use of the formula appears to decrease in the upper ranks. We also find a gender effect as shown by Figure 19: women use the formulae more often than men. When cross-tabulating gender with social rank, we establish a considerable gender difference in the upper ranks (Figure 20). A gender difference does not appear in the lower ranks, where the values for men are almost equal to those for women The L(M) C and U(M)C women pattern more or less alike, while U(M)C men use relatively few formulae. Despite the low absolute numbers, we may conclude that the low score for U(M)C men is in line with our expectations.
Figure 18. Frequency per 10,000 words, by social class of *dat weet God* “that knows God” in the seventeenth century.

Figure 19. Frequency per 10,000 words, by gender of *dat weet God* “that knows God” in the seventeenth century.

Figure 20. Frequency per 10,000 words, by social class and by gender of *dat weet God* “that knows God” in the seventeenth century.
In the eighteenth-century subcorpus, 9 tokens of the formula were found, only two of which are complete (4f, g).

(4) f. *dat wet god die kender hlle herten is*
   that knows God who expert all hearts is
   ‘that knows God, who is an expert of all hearts’

   g. *dat weet god die een kender der herten is*
   that knows God who an expert the+gen hearts is
   ‘that knows God, who is an expert of the hearts’

The other tokens are abbreviated variants such as (4h).

(4) h. *Dat Weet den almoegende God*
   that knows the almighty God
   ‘the almighty God knows that’

Numbers are too low to allow for quantification, but a few tendencies can be observed, which are in line with the patterns attested before. Despite the fact that the eighteenth-century subcorpus comprises a considerable number of letters linked to the U(M)C, viz. 261/384 or 68% of the subcorpus, only 4 occurrences are linked to the upper ranks. What is more, there are no tokens in the UC at all, and the 4 UMC tokens are all produced by female letter writers. The remaining 5 tokens are related to the L(M)C and produced by men as well as women.

### 3.6 Initiating discourse

The formula *ik laat u weten als dat* “I let you know that” (see Chapter 3, Section 6.2.1) is one of the most frequent formulae in the seventeenth-century subcorpus. It prototypically appears in the introductory passages of letters, following the greeting (5a). The formula offers a convenient way of ordering the discourse elements. The key word of the formula is the infinitive *weten* “know”, often combined with inflected forms of *laten* “let”. The subject is usually the first person singular form *ik* “I”, and the indirect object is *u* “you”, but there are other options, such as the first person plural (5d) or the third person singular as a subject.

The subordinator is usually *als dat* “lit. as that, that” (5a,b), sometimes simply *dat* “that” (5c,d). The formula may be interrupted, especially by the writer’s self-reference (5b) or by a form of address.

(5) a. *een vriendelycke groetensij sij gescreeven aen u l mijn seer en*
   a friendly greeting be written to you my very and
   wel *beminde man hans pieterse*
   well beloved husband Hans Pieterse
   *ick laet u l weten als dat ick en onse pieter nogh kloeck ende*
   I let you know as that I and our Pieter still strong and
gesont ben
healthy are
‘a friendly greeting be written to you, my very well beloved husband Hans Pieterse. I let you know that I and our Pieter are still in good health’

b. ick ul huijsvrou trintijs laet ul weten als dat
I your wife Trintijs let you know as that
‘I, your wife Trijntijs, let you know that [...]’

c. so lat ijck ul weten dat
so let I you know that
‘so I let you know that [...]’

d. wy u vader en moeder laaten u weten dat
we your father and mother let you know that
‘we, your father and mother, let you know that [...]’

The formula may also follow other introductory phrasings, especially deze dient “this serves” (5e) or deze is dienende (5f) (see Chapter 3, Sections 6.1.3, 6.2.1). In such cases, the formula is phrased as a subclause of purpose. The verb doen occasionally occurs as an alternative for laten (compare 5e and 5f).

(5) e. Deze Diendt om ue te Doen weten
this serves to you to do know
‘this one serves to let you know that’

f. Deser is diennende Om ue. Mijn Welstant ue te Laete Weete
this is serving to you my health you to let know
‘this one is serving to let you know my health/that I’m still healthy’

We also find tokens of ik laat u weten als dat “I let you know that” in other parts of letters besides the introductory part. In such cases, the formula is commonly preceded by voorts “further” or voorts soo “further so”, or variants thereof in the seventeenth century and verder(s) and its variants in the eighteenth century, marking the transition from one part of the discourse to another (cf. Chapter 3, Section 6.2.2, and below, Section 3.7). Only rarely do we find another connective introducing the following part of the discourse (5k,l).

(5) g. uort loet ijck ul ul wete als daet
further let I you you know as that
‘furthermore, I let you know that’

h. vorder laet ick u weten als dat
further let I you know as that
‘furthermore, I let you know that’

i. vort so lat ijck ul weten dat
further so let I you know that
‘furthermore, I let you know that’
j. *verders lat ik u Ed: weten als dat*  
   further let I you know as that  
   ‘furthermore, I let you know that’

k. *wijders laat ik u Ed: weten als dat*  
   further let I you know as that  
   ‘furthermore, I let you know that’

l. *nog laat ik uw weeten dat*  
   also let I you know that  
   ‘also, I let you know that’

Concentrating on the key words *weten* “know” and forms of *laten* “let” and *doen* “make, let”, we extracted all relevant instances of the formula from the seventeenth-century autograph subcorpus. This provided us with 138 tokens. Figure 21, which plots the frequency of the formula per 10,000 words across social class, again shows a very clear social distribution. The use of the formula decreases sharply as the social class index rises. In the LC, values peak at over 35 tokens per 10,000 words. In the UC, the total score is lower than five per 10,000 words. At the same time, the gender pattern we encountered before stands out (Figure 22) and it holds when cross-tabulating gender and social class (Figure 23).

Turning to the diachronic picture, we will, as before, first focus on the eighteenth-century results. In the eighteenth century, we still find many tokens of the formula, though fewer than in the previous century. In total, there are 94 tokens from the 1770s/1780s. The eighteenth-century results are given in Figures 24 and 25, presenting the distribution across social class and gender respectively.

![Figure 21](image-url)  
*Figure 21.* Frequency per 10,000 words, by social class of *ik laat u weten dat* “I let you know that” in the seventeenth century.
Figure 22. Frequency per 10,000 words, by gender of *ik laat u weten dat* “I let you know that” in the seventeenth century.

Figure 23. Frequency per 10,000 words, by social class and by gender of *ik laat u weten dat* “I let you know that” in the seventeenth century.

Figure 24. Frequency per 10,000 words, by social class of *ik laat u weten dat* “I let you know that” in the eighteenth century.
Figure 25. Frequency per 10,000 words, by gender of *ik laat u weten dat* “I let you know that” in the eighteenth century.

It is clear that the social distribution established for the seventeenth century (Figure 21) also holds for the eighteenth century (Figure 24): there is a sharp decrease towards the right-hand side, with the UC scoring no instances at all. What we furthermore find is a gender difference, but opposite to the familiar pattern. Figure 25 presents a reversed gender pattern in that men use this formula more frequently than women. This reversed pattern appears to hold for the LC and LMC when cross-tabulating social class and gender in Figure 26. This reversed gender pattern was also found in the L(M)C, in the case of the *van herten leed* formula (Section 3.4).

The diachronic picture is presented in Figure 27, which shows the total scores for both periods, taken from Figures 21 and 24.

Figure 26. Frequency per 10,000 words, by gender of *ik laat u weten dat* “I let you know that” in the eighteenth century.
Figure 27. Frequency per 10,000 words, by social class of \textit{ik laat u weten dat} “I let you know that” in the seventeenth and eighteenth centuries.

Figure 27 clearly shows the diachronic difference. Within each social class, we find a change towards fewer formulae, with the UC reaching zero formulae in the eighteenth century. Summing up, the formula \textit{ik laat u weten als dat} “I let you know that” is still frequent in the eighteenth century, though less frequent than in the seventeenth century. The diachronic difference is very clear in Figure 27, where the grey columns are consistently lower than the black columns, demonstrating the decrease in the use of the formula over time. Having examined no fewer than 130 letters from the eighteenth-century UC, we did not find a single attestation of the formula, neither in letters from men nor in letters from women. Therefore, we conclude that the change is led by the UC, and moreover, in the seventeenth century by UC men more than by UC women.

3.7 Continuing discourse

As discussed in Chapter 3, Section 6.2.2, \textit{voorts} “further” is a highly frequent formulaic expression for linking parts of the discourse: it signals discourse continuity as well as a topic shift. In the eighteenth century, it is often rendered \textit{verder} “further”. We will leave this lexical change aside, as the function and meaning of \textit{voorts} and \textit{verder} appear to be interchangeable. In Chapter 3 (Section 6.2.2), we presented many examples of variants of \textit{voorts}, and we also indicated its co-occurrence with expletive \textit{soo} “so”. We will not discuss possible functional or distributional differences between \textit{voorts} and \textit{voorts soo}, and merely observe that \textit{voorts} may be reinforced by \textit{soo}, without, in our view, any obvious semantic differences; compare (6a) and (6b). So we take \textit{voorts} and \textit{voorts soo} to be one type, which we refer to as \textit{voorts}. Regarding the eighteenth century, we will speak of \textit{verder}.
The topic shifting and discourse continuing short formula *voorts* “further” often introduces the formula *ik laat u weten als dat* “I let you know that” (6a, b), discussed in the previous section, or *is mijn schrijven als dat* “lit. is my writing as that” (6c, d), also discussed in the previous chapter (Section 6.2.2). It should be noted that we counted examples (6a, b) as tokens of *ik laat u weten als dat* “I let you know that” in the previous section, and here as examples of *voorts*. Apart from the combination with the two formulae mentioned, the expression *voorts* often also appears in the second half of letters, where it introduces greetings and wishes anticipating the closure of the letter (6e, f, g). Furthermore, there are cases where *voorts* functions as a sentence-initial topic shifter, and thus is not followed by another formula (6h, i, j). Finally, *voorts* is usually, though not always sentence initial (6k).

(6) a. *voorst* laet ick ul weten als dat
   further let I you know as that
   ‘furthermore, I let you know that’

b. *vort* so lat ick ul weten dat
   further so let I you know that
   ‘furthermore, I let you know that’

c. *vorder* is min scrivens dese als dat
   further is my writing this as that
   ‘furthermore, I am writing you that’

d. *vort* js mijn schrijue niet veel als dat
   further is my writing not much as that
   ‘furthermore, I am writing you nothing much, except that’

e. *voort* de groetenijs van alle vrijnden
   further the greeting from all friends
   ‘furthermore, greetings from all friends’

f. *vorder* de groetenenise aan sr heermans en zij huijs vrou
   further the greeting to Mr Heermans and his wife
   ‘furthermore, greetings to Mr Heermans and his wife’

g. *voorts* wensche wy vl al te samen duijsent goede nachten
   further wish we you all together thousand good nights
   ‘furthermore, we wish you all together a thousand good nights’

h. *voorts* den toe stant van ons vaderlant Is noch heel slecht
   further the situation of our fatherland is still very bad
   ‘furthermore, the situation of our country is still very bad’

i. *vorders* soo schrijft gij dat gij mijn wel wat gelt soudt stueren
   further so write you that you me well some money would send
   ‘furthermore, you write that you would send me some money’
j. vorder sal Ick Iou van onse reis wat schrijven
   further shall I you of our journey something write
   ‘furthermore, I shall write you something about our journey’

k. beware v vorder vor quaet geselschap
   beware you further of bad company
   ‘furthermore, beware of bad company’

We extracted all instances of *voort, voorts, vorder* etc. from the 219 letters in the seventeenth-century autograph subcorpus, which resulted in 180 tokens. We subsequently investigated the distribution of the formula across social class and gender; see Figures 28 and 29.

![Figure 28](image-url)

**Figure 28.** Frequency per 10,000 words, by social class of *voorts* “furthermore” in the seventeenth century.

![Figure 29](image-url)

**Figure 29.** Frequency per 10,000 words, by gender of *voorts* “furthermore” in the seventeenth century.
The results are largely in line with earlier findings. The black columns representing the total scores per social class drop steadily from 38 tokens per 10,000 words in the LC, to 24 in the LMC and to around 15 in the UMC and UC, and men use fewer formulae than women. When cross-tabulating gender and social class, this gender difference appears most strongly in the UMC, and to a lesser extent in the LMC, while the LC and UC show a reversed or an almost equal pattern (cf. Figure 30).

As before, we will explore the time variable by focusing on the eighteenth-century results first, and then turn to the diachronic picture. The eighteenth-century subcorpus comprises 267 tokens of *voorts* and *verder*, the latter variant being dominant. Figure 31 and Figure 32 present the eighteenth-century results, distributed across social class and gender respectively.

![Figure 30](image)

**Figure 30.** Frequency per 10,000 words, by social class and by gender of *voorts* “furthermore” in the seventeenth century.

![Figure 31](image)

**Figure 31.** Frequency per 10,000 words, by social class of *voorts* “furthermore” in the eighteenth century.
The social distribution in Figure 31 recalls the patterns attested above for most other formulae, from both the seventeenth and the eighteenth century. The frequency of the formula decreases from a fairly high value in the LC to a much lower value in the UC, where we find only a handful of tokens per 10,000 words. On many occasions we also encountered a gender pattern, with women producing more formulae than men within the social classes. This gender pattern was mainly found in the seventeenth-century subcorpus, and to a lesser extent in the eighteenth-century subcorpus. Here, we see the reversed pattern that we came across earlier: within the social classes, men produce more tokens of *verder* than women. Note, however, that the difference is fairly small, which is confirmed when cross-tabulating social class and gender (Figure 33).
Gender differences may occur, in both directions, the net effect of which being that no strong and stable gender differences appear, which is also clear from the results for the UMC and the UC.

Similar to the formula discussed in Section 3.6 (ik laat u weten als dat “I let you know that”) and contrary to many other formulae, verder “further” is still quite frequent in the eighteenth century (Figure 34). Figure 34 shows a small decrease in the LC, an increase in the LMC, a more or less stable frequency in the UMC and a considerable decrease in the UC. We should note that the relative proportion for the total subcorpora drops from 17.6 tokens per 10,000 words in the seventeenth century to 13.6 tokens per 10,000 words in the eighteenth century. The largest difference is attested in the UC, where numbers drop from 14.8 to 4.8 per 10,000 words.

3.8 Conclusions

In Section 2.4, we formulated hypotheses concerning the use and distribution of epistolary formulae in our seventeenth and eighteenth-century material. The main idea was that recourse to formulaic language must have been more prevalent among less-experienced writers. Since literacy, schooling and an orientation towards writing practice were socially stratified in terms of social class and gender, and moreover increased in the eighteenth century, we expected to find more formulae in letters from the lower classes than in letters from the upper classes, more in letters from women than in letters from men, and more in the seventeenth-century subcorpus than in the
eighteenth-century subcorpus. These expectations have been borne out by the results of the case studies in the preceding Sections 3.2 to 3.7.

Interestingly, almost all formulae studied gradually decrease in frequency of use from the left side of the figures to the right side, that is, from the LC to the UC. In many instances, within social classes women score consistently higher than men. In the eighteenth-century the gender differences level and occasionally we find a reversed pattern, viz. in the *van herten leed, ik laat u weten* and *voort* formulae. This reversed gender effect appears in the LC and LMC. What this indicates, in our view, is that eighteenth-century letters from the L(M)C are simply still very formulaic, whether they are written by men or by women. The gendered decrease in the use of formulae, then, becomes particularly evident in the UMC and in the UC, where, after all, it is men who in some areas score zero instances (see Figures above). The lowest frequencies are mostly produced by UC men. It should be remembered that important clues for assigning social classes were profession or occupation. The group of UC men typically consists of wealthy merchants, ship-owners, academics, commissioned officers and the like, that is, people who, if not well-educated, had a profession where extensive writing skills were needed. At this point in our analysis, we deem it safe to conclude that the orderly parallel distribution across the language community of writing experience and of formulaic language is a strong indication that writing experience also holds the explanation for the distribution of formulaic language. In Section 4, we will provide further evidence that will render this explanation even more probable.

4. **Variation and change as functions of writing experience**

In this section, we will discuss the social, gender and diachronic results that we found in relation to writing experience, and will deal with writing experience and historians’ ideas about the spread of literacy in the language community.

4.1 **The social literacy boundary**

Most formulae discussed in Section 3 occur often in the LC, gradually decrease in frequency in the LMC and the UMC, and appear least in the UC. Thus, the attested social variation is distributed in accordance with the social distribution of literacy and writing experience. The results provide evidence supporting the claim by Frijhoff & Spies (1999: 238) that a “real boundary” existed between less and more literate groups in society. Consider Table 6, where we list the frequencies per 10,000 words of all formulae discussed in Section 3, divided per social class. The first column indicates the century from which the data originate and the column “F/10,000 words” contains the overall average frequency of the formulae per 10,000 words, that is, in all social classes taken together.
Table 6. Frequencies per 10,000 words of all formulae under discussion, by social class, and the overall frequency per 10,000 words for the entire subcorpus.

<table>
<thead>
<tr>
<th>Both subcorpora, scores for men and women</th>
<th>LC</th>
<th>LMC</th>
<th>F/10,000 words</th>
<th>UMC</th>
<th>UC</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th Een vriendelijke groetenisse “a friendly greeting”</td>
<td>9.1</td>
<td>8.6</td>
<td>4.1</td>
<td>2.8</td>
<td>0.7</td>
</tr>
<tr>
<td>17th Kloek en gezond “strong and healthy”</td>
<td>12.7</td>
<td>6.4</td>
<td>5.1</td>
<td>4.1</td>
<td>4.4</td>
</tr>
<tr>
<td>18th Fris en gezond “fresh and healthy”</td>
<td>14.4</td>
<td>4.1</td>
<td>2</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>17th Van herte leed “very sorry”</td>
<td>9.1</td>
<td>7.7</td>
<td>5.4</td>
<td>5.1</td>
<td>1.5</td>
</tr>
<tr>
<td>18th Van herte leed “very sorry”</td>
<td>15.6</td>
<td>12.9</td>
<td>4.5</td>
<td>3.6</td>
<td>1</td>
</tr>
<tr>
<td>17th dat weet God “that knows God”</td>
<td>5.5</td>
<td>4.1</td>
<td>2.8</td>
<td>2.1</td>
<td>3</td>
</tr>
<tr>
<td>17th Ik laat u weten als dat “I let you know that”</td>
<td>36.4</td>
<td>23.2</td>
<td>13.5</td>
<td>10</td>
<td>4.4</td>
</tr>
<tr>
<td>18th Ik laat u weten als dat “I let you know that”</td>
<td>26.7</td>
<td>13.8</td>
<td>4.8</td>
<td>3.8</td>
<td>0</td>
</tr>
<tr>
<td>17th Voorts “further”</td>
<td>38.2</td>
<td>23.6</td>
<td>17.6</td>
<td>14.3</td>
<td>14.8</td>
</tr>
<tr>
<td>18th Verder “further”</td>
<td>34.4</td>
<td>30</td>
<td>13.6</td>
<td>14.8</td>
<td>4.8</td>
</tr>
</tbody>
</table>

As can be seen in Table 6, frequencies in the LC and the LMC are consistently higher than the overall average in the middle column, whereas frequencies in the UMC and the UC are consistently lower than the overall average. The two exceptions are in italics in the Table. In the case of the deviating 3 score for the UC in seventeenth-century *dat weet God* “that knows God”, we are dealing with the least frequent formula and previously we stressed the importance of taking caution here. The deviating UMC score in the eighteenth-century *verder* ‘further’ formula almost equals the seventeenth-century UMC score. Note that this formula was less clearly diachronically different, with the LC showing only a minor decrease, the LMC even an increase, and only a considerable UC decrease of 10 tokens per 10,000 words, namely from 14.8 to 4.8.

The frequencies of epistolary formulae in Table 6, which are mainly above average in the LC and the LMC, and below average in the UMC and the UC, are linguistic evidence supporting historians’ idea that there was a literacy boundary in Early Modern society. We should stress at this point that this does not mean that letters from the LC and the LMC are poorly written or hardly understandable. There is no reason to
assume that these letters have fulfilled their communicative functions less successfully than those from the UMC and the UC. The literacy boundary does not divide society into illiterate and unlettered groups on the one hand, and literate groups on the other hand, but into different groups according to their writing experience, in this case manifested by the average use of epistolary formulae.

Diachronically, it is interesting that historians consider rising literacy rates and increasing book production to be characteristic of the eighteenth century. In this period an increasing number of people are assumed to have become involved in public communication through the written word (via books, periodicals, pamphlets and so on), an involvement that intensified in the course of the century (Kloek & Mijnhardt 2001: 82). It has previously been stated in Section 2 that two-thirds of the male population and one-third of the female population were estimated to be able to write in the Northern Netherlands in the second half of the seventeenth century (Frijhoff & Spies 1999: 237), and that literacy rates increased to 80% of the male and 60% of the female population around 1800 (Kloek & Mijnhardt 2001: 81). In historical-sociolinguistic parlance, writing experience increased, both socially (because more people than before participated in the written culture) and exponentially (because their participation in the written culture intensified).

The results presented in the previous sections appear to corroborate these ideas linguistically. Moreover, when we add up the absolute numbers of all formulae (cf. Table 6) for each social rank in both periods, we can establish the average frequency per 10,000 words per social group per century. The results give an indication of the number of formulae one would expect to encounter in a piece of discourse of 10,000 words. Note that the numbers are minimum values as many letters contain more formulae than the six discussed in the previous sections and presented in Table 6. We computed such values for all social classes for both periods. Figure 35 visualises the results.

![Figure 35. Average frequency (per 10,000 words) of epistolary formulae, by social class and by period.](image)
As can be seen in Figure 35, in all social classes a sharp decrease over time appears. In the UC the largest decrease occurs: from 28.9 per 10,000 words to 6.6 per 10,000 words, i.e. a 77% decrease. In the UMC we find a decrease of 40% (from 38.4 to 23 per 10,000 words) and in the LMC and LC a decrease of 17% and 18% respectively (from 73.6 to 60.9 per 10,000 in the LMC, and from 110.9 to 91.9 in the LC). This ties in with historians’ ideas about the extension of the communication community generally, but also in particular with the common assumption that the widening of the written culture first and foremost affected the upper and middle classes (e.g. Kloek & Mijnhardt 2001). In the same vein, linguists have argued that, both in the Netherlands and in England, the eighteenth century saw an increase in middle-class concerns for the education of their children, and gave rise to a written culture where not, as before, the elite set the tone, but the middle classes just as much (Crowley 1996; Tieken-Boon van Ostade 2008: 104–106, 121; Beal 2004: 93–105; Rutten 2009b, 2012). If there is a link between writing experience and the use of formulaic language, then Figure 35 reflects the increased writing experience of the upper middle and upper classes, whereas the lower and lower middle classes remained relatively isolated from heavy involvement in the written culture.

4.2 Gender differences

The same approach was conducted for gender differences: for all formulae in Table 6 we added up the absolute numbers per gender and per period and established the total numbers per 10,000 words. Again the results in Figure 36 are minimum values on the basis of the six formulae under discussion. Both men and women show a decrease over time, although more strongly for women than for men: a decrease of 73% (from 71.5 to 19.3 per 10,000 words) against 30% (from 38.9 to 27.1 per 10,000 words). This change even results into a reversed gender pattern in the eighteenth century. What is

![Figure 36. Average frequency (per 10,000 words) of epistolary formulae, by gender and by period.](image)
most important, however, is that the remarkable gender difference in the seventeenth century is levelling in the eighteenth century. This completely matches with the observation made by historians that literacy differences (and what we call writing experience) are less prominent in the eighteenth than in the seventeenth century.

When cross-tabulating the results from Figures 35 and 36, we get a more detailed picture (Figure 37). In the UC, the seventeenth-century gender difference almost disappeared in the eighteenth century, when both men and women score ca. 6–7 formulae per 10,000 words. A similar development appears in the UMC: the remarkable decrease by women reduces the strong seventeenth-century gender difference to a situation where both genders score 21–27 formulae per 10,000 words in the eighteenth century. A similar reduction of gender difference is visible in the LMC, due to the decreasing use of formulae among women (the frequency among men is even somewhat higher in the eighteenth century). LC men show a small decrease which might lead us to conclude that the L(M)C frequency is relatively stable. LC women, however, again show a considerable decrease of formulae, similar to LMC women, and resulting into a clear gender difference in the LC in the eighteenth century. However, we have to be aware of the fact that the number of LC-letters written by women is extremely low, which prevents us from drawing too firm conclusions. A striking result of our analyses is that gender effects in general appear to be more pronounced in the seventeenth century than in the eighteenth century, especially in the LMC, the UMC and the UC. The levelling of writing practices among men and women appears to be largely due to both men and women progressively conforming to the writing conventions mainly associated with letters by UC men in the seventeenth century, that is, a style characterised by few formulae.
4.3 Incremental and levelling writing experience

In the seventeenth century, clear-cut gender differences were found mostly in the LMC, UMC and UC, while the LC scores for men and women were often quite similar (see Section 4.2). The case studies in Section 3 and the combination of social and gender variation as discussed in Sections 4.1 and 4.2 reveal a pattern of what we will refer to as *incremental writing experience*, where men score lower frequencies than women within the same social classes and women's scores are often similar to men's scores one social class down the social class index. In summary:

- Scores for UC women may roughly equal those for UMC men. See, for instance, *ik laat u weten als dat* “I let you know that” and *voorts* “further” in the seventeenth century.
- Scores for UMC women may roughly equal those by LMC men, see, for instance, *ik laat u weten als dat* “I let you know that” and *voorts* “further” in the seventeenth century.
- Scores for U(M)C women may roughly equal those for L(M)C men. See, for instance, *een vriendelijke groetenisse zij geschreven aan* “a friendly greeting be written to”, *kloek en gezond* “strong and healthy”, and *god de kender* “God the expert”, all in the seventeenth century.
- Scores for LMC women may roughly equal those by LC men, see, for instance, *ik laat u weten als dat* “I let you know that” in the seventeenth century.

Regarding the reduction of epistolary formulae, men are one social class ahead of women in these cases and, on the assumption that the use of formulae is related to writing experience, men thus appear to be one social class ahead of women in terms of writing experience. Towards the end of the eighteenth century, however, social and gender differences are gradually levelled out.

The ongoing increase in writing experience characterising the seventeenth and especially the eighteenth century levelled degrees of writing experience, in particular in the upper ranks, thereby partially depriving the phenomenon of writing experience from its incremental character. In the seventeenth century, similarly low scores for men as well as women were found in the UC, see for instance *kloek en gezond* “strong and healthy”, and *van herte leed* “very sorry”. In the eighteenth century, the number of formulae where U(M)C men and women share similarly low scores has risen, including, for instance, *fris en gezond* “fresh and healthy”, *van herte leed* “very sorry”, *ik laat u weten als dat* “I let you know that” and *verder* “further”.

In the UMC and the UC, therefore, letters by men and women become more similar, as women also reduce their use of formulae. The fact that writing practices in the UMC and the UC are levelling, for men as well as for women, indicates that schooling opportunities and the ongoing textualisation of society affected the writing experience of men as well as women. Nonetheless, we find strong distributional patterns when discussing the results in terms of the social class index. The UC and, to a lesser extent,
the UMC may experience levelling writing practices, but the “real boundary” dividing society into less literate and more literate people still holds in the 1770s/1780s (cf. Frijhoff & Spies 1999: 238).

4.4 Writing experience vs. social identity and style

In the introduction to the present chapter we mentioned another possible explanation for variation in formulaic language: the view of formulaic language as a group-related common practice, independent of the level of writing experience. Differences between social groups originating from relatively less and more writing experience may have become markers of group languages. If, for instance, less-experienced writers use more epistolary formulae, these formulae may become characteristic of their language and the language of the social groups with which they identify (cf. Laitinen & Nordlund 2012). This may happen even among experienced writers from these social groups, and even when writing experience increases over time. After having achieved the results in our case studies, we may argue that the orderly parallel distribution of social classes and epistolary formulae, with the steady decrease in frequencies towards the right side of the figures, makes such an interpretation in terms of social identities or group practices less probable. If identity roles or group practices were used as explanatory factors, an explanation would still be needed for why frequencies decrease over time, why we do not find the opposite pattern of high scores for the UC and low scores for the LC, why there are hardly any cases where men score consistently higher than women, why there are no formulae which are, for instance, typically LMC, or characteristic of UMC, et cetera. It would thus remain obscure why almost all formulae show the same pattern in terms of social class, and the same gender distribution as well. It would, in fact, be quite astonishing for all formulae to behave more or less alike.

The variational patterns found definitely favour an explanation in terms of writing experience rather than identity roles or group practices. Nevertheless, social identity roles may to some extent come into play as well. When comparing autographs to non-autographs in Chapter 5, for instance, a modest case can be made for linguistic practices associated with social identities or group practices.

In the previous sections, we already discussed diachronic differences in the use of formulaic language, as we did in Section 3, discussing individual formulae. Generally speaking, the use of epistolary formulae decreases between the 1660s/1670s and the 1770s/1780s. Some formulae vanished and some still appear in the eighteenth century, but to a lesser extent. The social and gender patterns found for the seventeenth century, with UC men commonly scoring the lowest frequencies, strongly suggest that the change in letter writing practices was led by UC men, and spread from there to other social classes and to the other sex. This is confirmed by the results for the eighteenth century where the lowest frequencies are still found for UC men, but UC women as well as UMC men and women are often quite close. In the LC and the LMC, formulae are still common, though less so than in the seventeenth century. We conclude therefore
that the changes in the use of formulaic language represent a classic case of change from above in its social interpretation.

The diachronic difference undoubtedly is a style difference and therefore the question arises of whether a difference in style can be considered an explanatory factor. Letter writing style changes from the seventeenth to the eighteenth century and style differences in letters and other text types have been attested for Early and Late Modern English as well. In letters a decreasing use of epistolary formulae was found, and other text types were shown to have developed more “oral” or involved characteristics such as the use of first and second person pronouns, private verbs (know, think), wh-questions, and that-deletion in subordinate clauses (Biber & Finegan 1989; Nevalainen 2001; Biber 2001). In our view, however, style differences cannot explain the social and gender distributions found. Moreover, the diachronic style difference itself would need to be explained.

It is an explanation in terms of writing experience that covers all variation and change found. Following Wray (2002) and Elspaß (2005), we consider formulaic language as a safe option for less-experienced writers who rely on lexicalised prefabs and thereby reduce the effort needed for writing. The social and gender distributions of literacy, of schooling opportunities, of the daily contact with written varieties, in other words, of writing experience, are reflected in the social and gender distributions of formulaic language. Diachronically, increasing schooling opportunities and rates of literacy, and the ongoing textualisation of society led to a decrease in the use of formulae. In other words, variation and change in formulaic language are considered as functions of writing experience. Following Kuiper & Haggo (1984), we may say that the fewer neophytes there are in writing, the more creative language is used, which implies fewer fixed formulae.

5. Conclusions and further discussion

In this chapter, we discussed social, gender and diachronic variation and change in the use of formulaic language. We may conclude that our three hypotheses (see 2.4) were corroborated. In general, formulae were more in use among women than among men, more in lower than in upper classes, and more in the seventeenth than in the eighteenth century. We explained the variation and change found with reference to parallel patterns in the distribution of writing experience and argued that fixed formulae could be seen as Formulierungshilfe “formulating help” for those who had less writing experience and/or less contact with written varieties of Dutch. The use of formulaic language definitely enabled less-experienced writers to reduce the processing effort of written communication. Higher rates of literacy, better schooling opportunities and educational careers, and more learned professions and occupations reduced the need for fixed formulae, more among men than among women, more in the upper classes than in the lower classes, and more in the eighteenth century than in the seventeenth century.
In Chapter 3, we described among others epistolary formulae from the eighteenth century, which, in the present chapter, were shown to decrease in use. Time and again, we found the same diachronic effect of disappearing formulae. Nonetheless, we wish to stress that we do not want to suggest that no or hardly any formulaic language occurred towards the end of the eighteenth century. As a matter of fact, text-constitutive formulae, notably text-type formulae, have been in existence for a very long time, and still exist. Present-day letter writers still use fixed formulae that are deemed necessary for the composition of a letter, and in order to mark the document in question as a letter. What has changed, however, is the use of extensive inter-subjective and Christian-ritual formulae, as discussed in this chapter. In Chapter 3, Section 2, we presented and discussed, by way of illustration, a typical seventeenth-century letter, made up of extensive formulae regarding greetings, health, contact, God and such like. Knowledge of these formulae enabled writers with little writing experience to nevertheless write a genuine letter, and these are exactly the formulae that disappeared. Furthermore, and more surprisingly, similar effects were found on text-constitutive formulae such as *ik laat u weten als dat* “I let you know that” and *voorts* and *verder* “further”. Expressions very similar to these still exist, and are still used in letter writing. Yet, it is striking that such sharp decreases in frequencies were attested. Clearly, the explicit and somewhat monotonous marking of textual elements through the use of these formulae eroded, and gave way to a more differentiated repertoire. Again, knowledge of just these formulae enabled writers with little writing experience to create coherence in their texts in a straightforward and generally accepted way.

Research into the social, gender and diachronic distribution of formulaic language gave us an important insight into the communicative strategies employed by lesser-skilled letter writers. Recourse to fixed formulae helped them overcome problems of formulation. This conclusion raises new questions about still other aspects of the letter-writing practice. What happened, for example, when senders were not lesser-skilled, but even unskilled (illiterate) or otherwise unable to write a letter? This requires a comparison of autograph letters with non-autographs, i.e. with letters written by either social or professional writers (cf. Chapter 1, Section 6). Furthermore, how did letter writers acquire the formulaic language that characterises these seventeenth- and eighteenth-century letters? Did they use letter-writing manuals? These questions will be addressed in Chapter 5.
Chapter 5

Detailing the writing process
Formulaic language, social and professional writers, and the influence of letter-writing manuals

1. Introduction

In Chapter 4, we demonstrated that for less-experienced writers formulaic language was a safe option in their letter writing by focusing on patterns of variation and change. Here we consider another aspect of the writing process: the writing of letters in the name of someone else. In this case of non-autographs, social or professional writers may also have used formulaic language as an appropriate tool or useful format, despite the fact that they should most likely be considered experienced writers. We distinguish autographs and non-autographs using a procedure developed by Nobels & van der Wal (2012; cf. Chapter 1, Section 6). As a consequence, we are able to compare autograph and non-autograph letters and to clarify the function of epistolary formulae in the particular case of writing in the name of others. Secondly, we will address the key question of how letter writers, be they (lesser) experienced, social or professional writers, acquired their formulaic language by examining the possible influence of letter-writing manuals. Both the comparison of autographs and non-autographs, and the role of letter-writing manuals yield theoretically important observations and conclusions and will lead us to fine-tune our general conclusions about formulaic language in our letter corpora.

After having introduced in Section 2 the non-autograph subcorpus used for the present chapter, we will report on a series of cases studies carried out on this subcorpus in Section 3, and compare the results with those obtained from the autograph subcorpus used in Chapter 4. Important differences will be discussed in Section 3.7, where we will also draw conclusions about the relationship between formulaic language and the autograph or non-autograph status of letters. In Section 4, we will get a more detailed view of the practice of writing in the name of someone else by focusing on a few specific examples of social and professional writers. Furthermore, we go into the question of why particular experienced writers would use formulaic language in their private letters (Section 5). In Section 6, dealing with our second key issue of the writing process, we will give a brief overview of contemporary Dutch letter-writing manuals and compare the epistolary formulae presented in these publications with the actual usage in our seventeenth- and eighteenth-century letters. The conclusions will be presented in Section 7.
2. The non-autograph subcorpus

Compiling our seventeenth-century corpus, the non-autograph or autograph status of letters was established by using our specially developed Leiden Identification Procedure (LIP; cf. Nobels & van der Wal 2012; Nobels 2013: 53–76). It is important to note that this procedure combines script and content analysis, and is not based on any linguistic feature (cf. Chapter 1, Section 6). In cases in which we had insufficient proof for either category, letters were ultimately labelled as having an ‘unclear’ status. For the present chapter we used only established non-autographs (and no letters of unclear status) in order to make the most convincing comparison with the subcorpus of autographs, used in Chapter 4. Table 1 presents the make-up of the non-autograph subcorpus, compared to the autograph subcorpus.¹

Table 1. The seventeenth-century autograph and non-autograph subcorpora.

<table>
<thead>
<tr>
<th></th>
<th>Number of letters</th>
<th>Number of words</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th c. autographs</td>
<td>219</td>
<td>102,000</td>
</tr>
<tr>
<td>17th c. non-autographs</td>
<td>116</td>
<td>45,000</td>
</tr>
</tbody>
</table>

3. Comparing autographs and non-autographs

In this section, we will compare the frequency of the six formulae discussed in Chapter 4 on the basis of the autograph subcorpus to their frequency in the non-autograph subcorpus introduced in the previous section. When the sender and writer are not one and the same person, the language of a letter cannot be related to the social characteristics of the identified sender. Since the non-autographs are thus not linked to a specific social rank or gender, we will compare the total frequencies of formulae per 10,000 words.

3.1 Greeting

The familiar seventeenth-century formula *een vriendelijke groetenisse zij geschreven aan* “a friendly greeting be written to” occurs frequently in the non-autograph subcorpus; see the following example (1).

¹. Due to the increasing rate of literacy, writer identification is less difficult in the eighteenth century compared to the seventeenth century. Examining the original sources, we found a high number of eighteenth-century autograph letters. Although we still found some non-autograph letters, these were excluded from our final eighteenth-century corpus that thus comprises only autographs (cf. Chapter 1, Section 6).
Table 2. Frequency of *een vriendelijke groetenisse zij geschreven aan* “a friendly greeting be written to” in the seventeenth-century autograph and non-autograph subcorpora.

<table>
<thead>
<tr>
<th></th>
<th>Number of tokens</th>
<th>Frequency/10,000 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th c. autographs</td>
<td>42</td>
<td>4.1</td>
</tr>
<tr>
<td>17th c. non-autographs</td>
<td>59</td>
<td>13.1</td>
</tr>
</tbody>
</table>

(1) *Een vriendelijke groetenisse sij geschreven aen mijn lieve beminde Man*

a friendly greeting be written to my dear beloved husband

Jan Jacobsz

’A friendly greeting be written to my dear and beloved husband Jan Jacobsz’

Focusing on the keyword *groetenisse*, we extracted all occurrences from the non-autograph subcorpus. The results are presented in Table 2, and compared with the results from the autograph subcorpus. As can be seen in Table 2, the non-autograph subcorpus, which is less than half as large as the autograph subcorpus, provides even more tokens, viz. 59. Normalised per 10,000 words, this means that the formula occurs 13.1 times, which is more than three times the frequency in the autograph subcorpus. The formula is usually the opening sentence of a letter. This means that whereas about 1 in every 5 autographs begins with this formula (42/219), even 1 in every 2 non-autographs (59/116) has the phrase *a friendly greeting*.

3.2 Communicating health

The formula *kloek en gezond* “strong and healthy, very healthy” is frequently used in the non-autograph subcorpus. We extracted all examples from the corpus focusing on the keyword *kloek*. As in Chapter 4, only the combination of *kloek en gezond* counted as an example, not the use of only *kloek* (2 tokens in the non-autographs), nor the use of variants such as *klouck en groot* “strong and tall” and *cluck en weluarent* “strong and prosperous”. The formula is almost always part of a full formulaic opening; see example (2).

(2) *vriendelycke groetenijssse aen myn seer bemiunde lijf lendert louverse*

friendly greeting to my very beloved love Lendert Louverse

*iék laet vl vete als dat wij altesaeme noch klack ende*

I let you know as that we all together still strong and

*ghesont sijn ende verhoepe dat het met vl mede sooe ijs*

healthy are and hope that it with you also so is

’friendly greeting to my very beloved love Lendert Louverse, I let you know that we are all still in good health, and I hope that you are also still healthy’

Table 3 presents the results for both the autographs and the non-autographs.
Table 3. Frequency of *kloek en gezond* “strong and healthy” in the seventeenth-century autograph and non-autograph subcorpora.

<table>
<thead>
<tr>
<th></th>
<th>Number of tokens</th>
<th>Frequency/10,000 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th c. autographs</td>
<td>52</td>
<td>5.1</td>
</tr>
<tr>
<td>17th c. non-autographs</td>
<td>72</td>
<td>16</td>
</tr>
</tbody>
</table>

As with a friendly greeting in Section 3.1, the formula *kloek en gezond* “strong and healthy” occurs more often in the non-autographs than in the autographs. The frequency per 10,000 words is 16, tripling the frequency in the autograph subcorpus. The absolute token count of 72 implies that on average, more than half of the 116 letters contain the present formula.

3.3 Communicating sorrow

The expression *het zou mij van herten leed wezen* “I would be very sorry to hear that” occurs frequently in the non-autograph subcorpus, and often in formulaic opening passages (3).

(3) *maer vaer heet anders  heet soude mij van haarten leedt  weesen dat*

  but were it different it would me from heart sorrow be that

  *weet godt almactig die een kenner aller harten is*

  knows God almighty who an expert all+gen hearts is

  ‘but if it were different, I would be very sorry to hear that. God almighty, who

  knows all the hearts, knows that’

We extracted all examples from the non-autograph subcorpus focusing on the keyword *leed* “sorrow”. The results are presented in Table 4. As in the previous sections, we conclude that the frequency of the formula in the non-autographs approximately triples its frequency in the autographs. The absolute number of 64 occurrences means that on average half of the 116 letters comprise this formula.

Table 4. Frequency of *van herten leed* “very sorry” in the seventeenth-century autograph and non-autograph subcorpora.

<table>
<thead>
<tr>
<th></th>
<th>Number of tokens</th>
<th>Frequency/10,000 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th c. autographs</td>
<td>55</td>
<td>5.4</td>
</tr>
<tr>
<td>17th c. non-autographs</td>
<td>64</td>
<td>14.2</td>
</tr>
</tbody>
</table>
Table 5. Frequency of *dat weet God* “that knows God” in the seventeenth-century autograph and non-autograph subcorpora.

<table>
<thead>
<tr>
<th></th>
<th>Number of tokens</th>
<th>Frequency/10,000 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th c. autographs</td>
<td>29</td>
<td>2.8</td>
</tr>
<tr>
<td>17th c. non-autographs</td>
<td>41</td>
<td>9.1</td>
</tr>
</tbody>
</table>

3.4 Communicating God’s omniscience

As the formulaic expressions discussed in the previous sections, the formula communicating God’s omniscience, exemplified in (4), is more frequent in the non-autographs subcorpus than in the autographs subcorpus; see Table 5.

(4) *dat weet Godt al machtich die een kender is van aller menschen hert*  
that knows God almighty who an expert is of all people heart  
‘God almighty, who knows the hearts of all people, knows that’

With 9.1 tokens per 10,000 words, the formula occurs more than three times as frequently in the non-autographs as in the autographs.

3.5 Initiating discourse

The performative formula *ik laat u weten dat* “I let you know that” was highly frequent in the autograph subcorpus, occurring 138 times in 219 private letters, which amounted to an average of 13.5 tokens per 10,000 words. As in the autographs, in the non-autograph subcorpus the formula is sometimes part of formulaic opening passages (5), but it also appears further down in letters, introducing a new topic (6).

(5) *jck joesghen korneles ul bemijnde huijsvrou late ul weten als dat jck noch kloeck ende gesont ben*  
I Joesghen Korneles your beloved wife let you know as that I still strong and healthy am  
‘I, Joesghen Korneles, your beloved wife, let you know that I am still in good health’

(6) *uoors soo laet ick u min beminde man weeten als dat u suster meinsie haer kint aen de pest gelegen heit*  
Meinsie her child on the plague lain has  
‘furthermore, I let you, my beloved husband, know that your sister Meinsie’s child has suffered from the plague’

Focusing on the keyword *weten* “know”, we extracted all tokens from the non-autograph subcorpus, comparing these to the results of the autograph subcorpus in Table 6.
Table 6. Frequency of *ik laat u weten dat* “I let you know that” in the seventeenth-century autograph and non-autograph subcorpora

<table>
<thead>
<tr>
<th></th>
<th>Number of tokens</th>
<th>Frequency/10,000 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th c. autographs</td>
<td>138</td>
<td>13.5</td>
</tr>
<tr>
<td>17th c. non-autographs</td>
<td>148</td>
<td>32.9</td>
</tr>
</tbody>
</table>

Table 6 shows that the formula *I let you know that* is even more frequent in the non-autographs than in the autographs, occurring 148 times in just 116 letters. Per 10,000 words, the frequency amounts to a considerable 32.9.

3.6 Continuing discourse

The most frequent formula in the autograph subcorpus was the discourse marking expression *voorts (so)* “further (so), furthermore”, which signals the continuation of the discourse. It can be followed by another formulaic expression as in (6) above and in (7), but also by non-formulaic, context-dependent information (8).

(7) *voort laat ick ul weeten als dat ghij dient te weeten hoe desoolaet dat ons lieve vaderlandt staet* further let I you know as that you need to know how desolate that our dear fatherland stands

‘furthermore, I let you know that you need to know how desolate our dear country is’

(8) *voorts soo hebbe ick alle daghen van u moeder wennighe vrintschap* further so have I all days from your mother little friendship

‘furthermore, I usually enjoy little friendship from your mother’

Extraction of all relevant examples from the non-autograph subcorpus provides the results presented in Table 7. As can be seen in Table 7, the formula *voorts* “furthermore” is also highly frequent, and in fact the most frequent formula, in the non-autographs subcorpus, occurring 166 times in 116 letters. Compared to the autograph subcorpus, its frequency rises from 17.6 to 36.9 per 10,000 words.

Table 7. Frequency of *voorts* “furthermore” in the seventeenth-century autograph and non-autograph subcorpora.

<table>
<thead>
<tr>
<th></th>
<th>Number of tokens</th>
<th>Frequency/10,000 words</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th c. autographs</td>
<td>180</td>
<td>17.6</td>
</tr>
<tr>
<td>17th c. non-autographs</td>
<td>166</td>
<td>36.9</td>
</tr>
</tbody>
</table>
Chapter 5. Detailing the writing process

Furthermore

I let you know

God knows

Strong and healthy

A friendly greeting

Figure 1. Frequency of six formulae in the seventeenth-century autograph and non-autograph subcorpora.

3.7 Conclusion

In the previous sections, striking, if not spectacular frequency differences in the use of six epistolary formulae have been established when comparing the results for the autograph subcorpus, as presented in Chapter 4, to the results of exactly similar searches in the non-autograph subcorpus. Figure 1 summarises these results.

As Figure 1 shows, the frequency per 10,000 words of the four formulae ‘een vriendelijke groetensij geschreven aan’ “a friendly greeting be written to”, ‘kloek en gezond’ “strong and healthy”, ‘het zou mij van harte leed zijn om te horen’ “I would be very sorry to hear that”, and ‘dat weet God die een kenner is van alle harten’ “God who is an expert of all the hearts knows that” approximately tripled. Their frequencies in the seventeenth-century autograph subcorpus are all in the range from 2.8 to 5.4 against the range from 9.1 to 16 in the seventeenth-century non-autograph subcorpus. The frequency of the formula ‘ik laat u weten dat’ “I let you know that” multiplied by 2.5, rising from 13.5 to 32.9 per 10,000 words, while the formula ‘voorts’ “furthermore” doubled its frequency, from 17.6 in the autograph subcorpus to 36.9 in the non-autographs. Thus the letters in the non-autograph subcorpus appear to be more formulaic than the letters in the autograph subcorpus. The frequency differences established in the previous sections and summarised in Figure 1 constitute a stable pattern, clearly suggesting that letters written in the name of someone else are in general much more formulaic than autographs. Linking this to the writing process, we suspect that whenever someone with little or no writing experience relies on a social or a professional writer, these writers will be inclined to fall back on formulaic
language. The data suggest that social and professional writers must have had recourse to formulaic language, but self-evidently not because of a lack of writing experience, since these writers were, on the contrary, called upon for their expertise. Another important factor must therefore be at play in the distribution of formulaic language in non-autographs: not the lack of writing experience, but fixed formatting, convenient to social and professional writers.

We may conclude that formulaic language was not just a safe option for less experienced writers, it was also a safe option for social and professional writers. This seems to be an obvious idea. Without further evidence, however, it is simply a plausible hypothesis. Here we have presented quantitative linguistic evidence which proves that it must have been common practice for social or professional writers to rely on formulaic language. Often we do not know who the social or professional writers were, but in some cases we discovered their identity. In the following section we will present a few particular examples and show where the senders looked for writing assistance.

4. Writing in the name of: Social or professional writers

Various examples show that people looked for writing assistance, not only in the same region, but also in their immediate social environment, both in terms of gender and in terms of social rank. Neeltje Jacobs, a UMC woman, of whom we have both an autograph and a non-autograph letter, is such an example. Having just given birth in 1672, she asked Trijntje Cornelis, another woman categorised as UMC, to write a letter for her. Still, there are instances where gender or social boundaries were crossed. Joosje Cornelis from the UMC, for instance, sent a letter to her husband in 1664. We identified the letter writer who was Kathelijne Haexwant, a woman one step higher on the social class index, viz. belonging to the UC. Similarly, Helena Backx from the LMC sent a letter to her brother Christiaan in 1672, which was written by another woman (Elisabeth Emerij), whom we categorised as UMC. The letter by Dorothee Pieters from the LC is an example of crossing the gender boundary. In 1664, she sent her husband a letter, which was probably written by her brother Willem Pieters. These examples demonstrate that, when social or gender boundaries were crossed, this unsurprisingly happened upwardly in the direction of more writing experience according to the social and gender distribution of writing experience established in Chapter 4.

By way of illustration, we present one of these letters as an insightful example: the letter sent by Joosje Cornelis (or Joesghen Korneles as written in example (9)), a woman from the UMC, whose 1664 letter was written by Katelijne Haexwant from the UC.
Figure 2. Letter, dated 31 October 1664, sent by Joosje Cornelis and written by Kathelijne Haexwant.

After the address formula, the letter runs as follows.

(9) vriendelijke groetenissen aen ul mijn bemijnde man jonas pietersen
friendly greeting to you my beloved husband Jonas Pietersen
jck joesghen korneles ul bemijnde huijsvrou late ul weten als dat jck
I Joesghen Korneles your beloved wife let you know as that I
noch kloeck ende gesont ben met alle onse kijnderen godt lof voor
still strong and healthy am with all our children God praise for
sijn groete genade verhoepende dat het met ul mijn lieue man
his great mercy hoping that it with you my beloved husband
oock soo is ware het anders ’t soude mijn van harten leet sijn dat
also such is were it different it would me from heart sorrow be that
weet godt almachtijch die een kender van alle harten is voors soo
knows God almighty who an expert of all hearts is further so
doet maerghen tyssen haer soon jan arijensen van harten groeten en
do Maerghen Tyssen her son Jan Arijensen from heart greeting and
sij is noch klock en ul moeder oock sij hadde mijn geseijt als
she is still strong and your mother too she had me told if
jck eens schref dat jck eens aen ul soude schrijuen sij en weet
I once wrote that I once to you should write she neg knows
het nou niet jck en dorst het haer niet seggen voors soo is onse katerijna
it now not I neg dare it her not tell further so is our Katerijna
stoeps noch klock en sij doet haer noom groeten en jck ul
Stoeps still strong and she does her uncle greeting and I your
meueghen doen ul oock groeten en wensschen ul altesamen
aunt do you also greeting and wish you all-together
hondertduijsent goeden nacht en een gelucksalijghe reijse dat wij
hundred-thousand good night and a blessed journey that we
malkanderen moghen sien en spreken amen den 31 oktober 1664 bij
each-other may see and speak amen the 31 october 1664 by
mijn joesghen korneles ul hujsvrou wat jck vermach tot ul
me Joesghen Korneles your wife what I can to your
besten amen
best amen
‘A friendly greeting to you, my beloved husband Jonas Pietersen. I, Joesghen
Korneles, your beloved wife, let you know that I am still strong and healthy
with all our children. Praise God for his great mercy. I hope that you, my dear
husband, are also like that. If it were different, I would very much regret to
have to hear that. That knows God almighty, who knows all the hearts.
Furthermore, Maerghen Tyssen sends hearty greetings to her son Jan Arijensen.
And she is still healthy, and your mother too. She had told me, that if I would
write again, that I would write to you. She does not know it now. I did not dare
to tell her. Furthermore, our Katerijna Stoeps is still healthy, and she sends her
uncle her greetings, and I, your aunt, send you also my greetings, and I wish
you all together a hundred thousand good nights and a blessed journey. That
we may see and speak each other. Amen. 31 October 1664. By me, Joesghen
Korneles, your wife. Whatever I am capable of to do in your interest. Amen’

We have marked the formulaic parts in boldface. The letter opens with an almost com-
plete paragraph with greeting, health and Christian-ritual formulae, followed by voors
soo “further”, marking the transition to the less formulaic middle part, which however
starts off with greetings and health statements. A very brief non-formulaic interlude
follows, which is then cut off with another voors soo, introducing greetings and health
formulae again. Then the letter comes to an end with wishes, contact formulae, the
date, the name of the sender (not the writer), interwoven with closing formulae such
as amen and wat jck vermach tot ul besten “whatever I am capable of to do in your in-
terest”. With this letter, the sender Joosje Cornelis and the writer Katelijne Haexwant
demonstrate the form that everyday communication may ultimately take.

Furthermore, people who were unable to write a letter were sometimes assumed
to have called upon professional writers (Métayer 2000). So far, we have found very
little concrete evidence of an important role for professional writers such as the writers
on board ships and notaries or school teachers ashore.2 In the relatively few cases
where the writer of a non-autograph can be identified, we are almost always dealing

2. See Braunius (1964) for the tasks of writers on board ships.
Figure 3. Letter, dated 6 August 1664.
Figure 4. Letter, dated 31 October 1664.
with a so-called social writer, that is, a person from the sender’s social environment. In the highly urbanised Northern Netherlands where literacy rates were remarkably high (Frijhoff & Spies 1999: 237), finding such a social writer could not have been very difficult. Note that people not always called upon the same person to write a letter for them: Hendrikje ten Broek, for instance, an eighteenth-century UMC woman, sent her husband non-autograph letters in two different hands which we were not able to identify. In the case of Geertruyt Weckmans, a seventeenth-century LMC woman, however, we discovered the identities of the helpful letter writers who, most interestingly, appeared to be a social and a professional writer, respectively. In 1664, Geertruyt Weckmans sent two letters to her former mistress Geertruyt Boetselaer in different handwritings, signed with her name in yet another (her own) handwriting.

Both letters begin with the familiar seventeenth-century formula expressing praise to God (cf. Chapter 3, Section 6.1.2): the letter dated 6 August 1664 (Figure 3) with the Latinized phrase *Laus Deo Semper* “praise God always” and the letter dated 12 October 1664 (Figure 4) with the more common Dutch variant *Looft godt bouen all* “praise God above all”. An analysis of the two letters revealed a number of striking differences and similarities which will not be discussed here (for the analysis cf. van der Wal & Rutten 2013a). What these exemplary letters show is what kind of persons functioned as social or professional writers. The first letter appeared to be a rare case of an identified professional writer: the secretary and part-time school master Ludovicus Cobes/Cobus. The second letter was written by a social writer from Geertruyt Weckmans’ immediate and close environment: her husband Pieter Meessen Vroman (van der Wal & Rutten 2013a: 129, 135).

The examples discussed in this section allow us to see a glimpse of the actual practice of calling upon more experienced letter writers. Before turning to the question of how both writers of autographs and social or professional writers acquired the formulaic language that characterises many seventeenth- and eighteenth-century letters, we once more present an insightful example to address a remaining question about the use of formulae.

### 5. Formulaic language, group practices and social routine

While we gained insight into the social and gender distribution of formulaic language on the one hand (Chapter 4) and the strikingly high frequency of formulae in non-autographs on the other hand (Section 3), we still have to deal with a few remarkable examples of frequent formulaic language used by particular experienced writers in their own private letters. Katelijne Haexwant from the UC is such an experienced writer who uses formulaic language very frequently. But her letters are not simply very

---

3. Cf. van der Wal, Rutten & Simons (2012: 144–146). Hendrikje ten Broek’s non-autograph letters were not included in our eighteenth-century corpus which consists of only autographs.

4. We thank Janny Venema and Charles Gehring (New Netherland Research Center, Albany) for their identification of Ludovicus Cobus’ handwriting.
formulaic, they are also long, and loaded with information. The following example (10) is the beginning of a letter sent to her husband Leendert Arijensen Haeswant on 31 October 1664, with its 146 lines and 1952 words, an exceptionally long one. Recall from Chapters 2 and 4 that seventeenth-century letters comprise 450 to 470 words on average. Although its length is a clear indication of the writer’s well-developed writing skills, the letter begins with a ten-line full formulaic opening.

(10) **Looft Godt In rotterdaem den 31 oktober 1664**  
Praise God in Rotterdam the 31 October 1664  
Vriendelijcke Groetenisse aen vl Mijn lieue ende wel bemijnde man  
friendly greeting to you my dear and well beloved husband  
Leendert arijensen haeswant Ick katelijngen haeswants vl bemijnde  
Leendert Arijensen Haeswant I Katelijngen Haeswants your beloved  
huijsvrou late vl weten als dat Ick noch kloeck ende gesont ben godt  
wife let you know as that I still strong and healthy am God  
lof verhoepende dat het met vl mijn lieue ende bemijnde man  
praise hoping that it with you my dear and beloved husband  
oock soo is ware het anders tsoude mijn van harten leet sijn  
also such is were it different it-would my from heart sorrow be  
dat weet godt almachtijch die een kender van alle harten is voors  
that knows God almighty who an expert of all hearts is further  
soo doen alle onse lieue kijnderen vl groeten en broers en susters  
so do all our sweet children you greet and brothers and sisters  
oock en sjn altesamen noch klock godt dank voors soo late Ick  
too and they are altogether still strong God thank further so let I  
vl weten dat Ick een brief ontfanghen hebbe van den 2 september een  
you know that I a letter received have of the 2 September one  
vanden 10 een van den 20 en van den 28 alle van september een  
of-the 10 one of the 20 and of the 28 all of September one  
vanden 1 oktober met een schepghen op seelant ende den in hout  
of-the 1 October with a little ship to Zeeland and the contents  
wel verstaen hebbe ...  
well understood have ...  
‘Praise God. Rotterdam, 31 October 1664  
Friendly greeting to you, my dear and well beloved husband Leendert Arijensen Haeswant. I, Katelijngen Haeswants, your beloved wife, let you know that I am still strong and healthy. Praise God. Hoping that you are also like that, my dear and beloved husband. If it were different, I would very much regret to have to hear that. That knows God almighty, who knows all hearts. Furthermore, all our sweet children send you their greetings, as well as the brothers and sisters. And they are all together still healthy. Thank God. Furthermore, I let you know that I received a letter dated 2 September, one dated 10 and one dated 20 and [another] dated 28, all dated September, one dated 1 October with a little ship to Zeeland. And I have understood the contents ...’
Clearly, if writing experience explains the distribution of epistolary formulae in the language community, as argued in Chapter 4, the abundant use of formulae by an experienced writer such as Katelijne Haexwant, who even functioned as a social writer for others, remains unexplained. In order to find an explanation, we again discuss the possibility that formulaic language was not only related to writing experience, but also to writers’ identities, social roles or group practices (see also Chapter 4, Section 4.4).

It has been pointed out before that linguistic choices, including epistolary formulae may index writers’ personal or social identities (Pahta et al. 2010: 1–6; Wood 2004; Laitinen & Nordlund 2012). Although we will not rule out the possibility that a different, more qualitatively oriented investigation of the letters might reveal significant patterns in the distribution of formulae in connection with social roles or identities, we argue as before (Chapter 4, Section 4.4) that the recurrent social and gender patterns certainly do not favour an interpretation in terms of social identity. It is all too telling that so many formulae pattern so similarly according to social class, gender and time. Regardless of the history of epistolary formulae in well-defined speech acts in official texts (Davis 1965; Nevalainen 2001), the well-ordered distribution of epistolary formulae in the seventeenth-century language community of the Northern Netherlands favours an interpretation in terms of writing experience. However, in the case of non-autographs, written by social or professional writers, the fixed formatting explanation was shown to apply and in a few particular cases, such as Kathelijne Haexwant, a social identity explanation and familiarity with epistolary conventions can be considered. Taking into account the latter explanation, we have to realise that, as highly frequent elements of epistolary language, the formulae were conventions. These conventions may develop into social routine formulae suitable for standardised communicative situations (cf. Coulmas 1981a). In other words, the formulae lingered on in usage as conventions, having lost not just their original pragmatic function, but also the connection to writing experience typical of seventeenth-century usage. Therefore, we may understand that experienced writers such as Katelijne Haexwant could maintain formulae in their usage, thus demonstrating their familiarity with epistolary conventions.

After having gained insight into the actual writing process and into the differences between autograph writers and social or professional writers, we turn to the second main issue of this chapter, that is the frequently arising question of how letter writers acquired formulaic language and what role letter-writing manuals played. To answer this question, we will examine formulae in contemporary letter-writing manuals and compare these with the extensive inventory of formulae discussed in Chapter 3.

6. Formulae in letter-writing manuals

So far we have focused on formulae in actual language use, without considering the vast body of letter-writing manuals, the earliest dating back to the sixteenth century, and many of them passing through several reprints in the course of the seventeenth and eighteenth centuries. The Dutch situation was very similar to that in England, where the
eighteenth century in particular witnessed a strong increase in the number of letter-writing manuals (Bannet 2005; Mitchell 2007). The increased production naturally raises the question of how the manuals functioned in letter writers’ daily life, and to what extent they influenced letter-writing practice. In this section, we will, therefore, briefly discuss Dutch letter-writing manuals, compare the epistolary formulae presented in these publications with the actual usage in our seventeenth- and eighteenth-century letters and reflect on possible influences of epistolary norms on actual usage, and vice versa.

The reader should keep in mind that our corpus data are from the 1660s–1670s and from the 1770s–1780s, respectively, but that we encounter a long-standing tradition of meta-epistolary comment. Most seventeenth-century letter-writing manuals date back to the late sixteenth or early seventeenth centuries, with efforts at renewal surfacing in the course of the eighteenth century. Furthermore, relying on title pages, prefaces, contemporary references to the letter-writing manuals and other bibliographical information, a distinction can be made between manuals apparently aiming at an audience fairly high up the social scale, and manuals targeted towards a more modest audience and/or intended to be used in schools. We will refer to the former as elite manuals and to the latter as school books, without, however, claiming that the elite manuals were never used in school, or that the school books were never used among the upper classes.

6.1 Elite manuals

An early well-known letter-writing manual is the Nederduytse secretaris oft zendbriefschryver “Dutch secretary or letter writer” (1637) by Daniël Mostart, an author who associated with literary figures such as the poets P.C. Hooft and Joost van den Vondel, and who was, apart from his profession of secretary, a playwright himself (Koning 1997). As secretary of the city of Amsterdam, Mostart dedicated his manual to the bailiff, burgomasters and aldermen of Amsterdam. Comprising over 300 pages, the manual is first and foremost a companion to the secretary’s profession, incorporating both extensive instructions as well as ample example letters, many of which were addressed to rulers and governments. It is therefore not surprising that the formulae from our letter corpora do not appear with any frequency in this manual. One example is found in the closing of a letter on page 197, where letters between friends are discussed:

(11) Ik verlang uwer Ed: gezontheit te vernemen, met my is ‘t, God I long your honour’s health to understand, with me is it God lof wel. Hier mede. praise well here with ‘I long to understand your honour’s health. With me, all is well, praise God. With this.’

This phrase is reminiscent of the formulaic practices described in the present and the previous chapters, albeit that health statements are usually found in the introductions of letters (see Chapter 3, Section 7.1). Note also the abbreviated formula Hier mede “lit. here with, with this”, used for closing discourse (see Chapter 3, Section 6.2.3). This closing
formula appears throughout Mostart's manual, and it is in fact the most obvious parallel between his model letters and the actual letters discussed in the present monograph.

Similarly, the Fatsoenlycke send- brief- schryver “Decent letter writer” (1654) by de la Serre, translated from de la Serre’s French original, was intended for society’s higher echelons. It contains fifty pages of instruction, three hundred pages with model letters, and another one hundred pages with rules of etiquette and such like. The model letters are not formulaic, or at least do not contain many formulae found in our letter corpora. One salient parallel with many of our eighteenth-century letters is the repetition of the addressing of the recipient just before closing the letter (see Chapter 3, Section 6.1.4). This may have been an upper-class convention in the seventeenth century, which spread through society in the subsequent decades. The convention is, for instance, found throughout the correspondence of P.C. Hooft, both in his own letters and in letters addressed to him (van Tricht 1976–1979). It is also discussed in an eighteenth-century successor to Mostart and de la Serre, viz. Le nouveau et parfait secretaire en hollandois et francois/De nieuwe volmaakte Hollandsche en Fransche secretaris “The new and perfect Dutch and French secretary”5 This is once more a very extensive guidebook, the first eighty pages containing instructions in Dutch, then some eighty pages with the same instructions in French, and finally about 450 pages with example letters in both Dutch and French. On page 59, it is explained that the vocativus has to be repeated when closing a letter. A discrepancy between this eighteenth-century manual and actual practice is the advice to place the date at the top of the letter, which contradicts the development in our corpora (see Chapter 3, Section 6.1.2). As with Mostart and de la Serre, the many example letters hardly contain any formulae found in our letter corpora.

Less voluminous, but with its 183 pages still quite sizeable is the Nieuwe handleiding tot de manier van brieven schryven “New manual to the art of letter writing”6 This is also an eighteenth-century manual with a fairly detailed instruction and many example letters. Again, hardly any of the expressions found in the book match the epistolary formulae commonly found in our letter corpora. What we do find, however, are rare examples of the extensive closing formulae that we encounter in the eighteenth-century part of our corpus (see Chapter 3, Section 6.1.4). See examples (12, 13) from this Nieuwe handleiding “New manual”.

(12) Bedanken UWelEd. voor de eere der Communicatie, en thank you for the honour of communication, and blyven naar my in UEd. waarde Vriendschap aanbevolen te remain after me in your worthy friendship commended to hebben, met alle achting. have with all respect ‘[I] thank you for the honour of communication, and [I] remain with all respect, after having me commended into your friendship’

5. We used the second, 1750 edition, published in Rotterdam.
6. We used the third edition, published in Amsterdam in 1770.
It appears that these extensive closing formulae were an element in both the normative tradition and in actual practice. Note that in (13), we also encounter the phrase *naar vriendelyke groetenisse* “after friendly greeting”, which we also found in our letter corpora (Chapter 3, Section 6.1.3), albeit mostly in the openings of letters.

The manuals discussed so far probably aimed at a readership fairly high up the social scale. This is clear from their size, from the example letters addressed to rulers and governments, and from the incorporation of French. Similarly, Breton’s *Nederduytsen briefdragher* “Dutch postman” (1645), translated from English but originally a French work, is of no real importance for the interpretation of the language use in our letter corpora. After all, Breton’s work is part of “an ongoing tradition of ‘Merriment’ or witty take-offs on standard letters” (Bannet 2005: xv). In the next section, we will discuss manuals that may have been of more importance to the letter writers in our corpus.

### 6.2 School books

There were also writing manuals with a more modest target audience and/or that were oriented towards schools. It is in these manuals that we find many more resemblances to formulaic language as it is found in the actual practice of our letter writers. Examples from the late sixteenth and the first half of the seventeenth century are the manuals by Jacobi, Smyters and de la Chambre, as well as the *Matery-boecksken* “book of matter”. In the late seventeenth and in the eighteenth centuries, authors Hakvoord and Koolenkamp aimed at replacing Jacobi’s manual – Koolenkamp mainly by modernising it – but Jacobi’s manual was still published throughout the eighteenth century, and even early in the nineteenth century. We will discuss these books in more detail, indicating the similarities to and differences from the actual language use.

Both Anthoni Smyters and Pierre de la Chambre were Holland-based school-teachers, who wrote a letter-writing manual explicitly targeted at the young pupils in their schools. Smyters, who originated from Antwerp but taught in Amsterdam at the time, dedicated his *Schryf-kunst-boeck* “lit. writing-art-book, book about the art of writing” (1613) to the Amsterdam school inspectorate, explaining that it may be used as a textbook. The book begins with instructions about the art of letter-writing and orthographical guidelines (Dibbets 1986). The main part (pp. 26–109) is devoted to 105 model letters, with a few business letters and receipts. The book gives the impression of having emanated from hand-written examples used by Smyters. By printing his
teaching materials, he took the opportunity of making some profit from it. The same applies to de la Chambre’s book, which will be discussed below.

In both Smyters’ instructions and in the model letters, we encounter phrases that resemble the formulae found in actual usage. Adjectives suitable for characterising the addressee include *eersamen* “honest” and *zeer beminden* “very beloved” and the particle *woonende* “living” occurs in the address (1613: 20). A discourse continuing formula as well as the commendation formula occur in an example such as *voorts den almachtigen bevolen* “further commended to the Almighty” (1613: 21). Consider example (14), where we give the first part of an example letter (1613: 56):

(14) Nae vriendelijke groete/ dienen dese om u.l. te adverteren/ dat ick God after friendly greeting serve this to you to inform that I God *lof wel te passe ben*. Verhopende ende van herten wenschende sulckx praise well in order am hoping and from heart wishing dat *met u.l. te wesen*. Ten ander dat ick met blyschap uwe Brieven with you to be at-the other that I with happiness your letters *ontfanghen hebbe/ende daer uyt verstaen uwe goede* received have and there from understood your good *ghesondtheydt / t’welck my seer lief om hooren is gheweest*: Waer health the-which me very dear to hear is been where *voren wy God behooren te dancken: Ende hem bidden / dat het hem* for we God should to thank and him pray  that it him *ghelieve ons die langhe te laten ghenieten? tot onse saligheyt*. Voorder... please us that long to let enjoy until our salvation further ‘After friendly greeting, this [letter] serves to inform you that I am in good health, praise God. Hoping and wishing from the heart that you are also in good health. Secondly, that I have received your letters with happiness, and I have understood from them your good health, which I was very pleased to hear, and for which we have to thank God, and pray to Him that he may let us enjoy it for a long time, until our salvation. Furthermore...’

Many of these formulae appear elsewhere in Smyters’ book as well as in our letter corpus. More examples of resemblances could be given, but what should also be mentioned is the presence of epistolary formulae which do not occur in our letters, for instance *d’Oorsake mijns schryvens is* “the reason of my writing is” (1613: 55), and opening phrases such as *Ick recommandere my seer aen U.L.* and *Ick ghebied my zeer dienstelijken t’uwaerts* “I very much (re)command myself to you” (1613: 96, 101). On the other hand, there are also formulae from actual usage which are not found in Smyters’ manual, e.g. the formula *kloek en gezond*, where Smyters consistently uses *wel te passe* “well to state, healthy”.

Pierre de la Chambre, a schoolteacher in Beverwijk (North Holland), published his bilingual manual *Verscheyden brieven, bequaem in de scholen te ghebryucken ...*
Plusieurs lettres missives, a l’usage des écoles “Various letters, suitable for use in the schools” in Haarlem in 1648. As he explains in his one-page foreword (page A1’), he composed his manual first and foremost “for the profit and use of my own school”. The manual contains 36 letters, all of them presented both in Dutch (on the left-hand page) and in French (on the right-hand page). As with Smyters, there are striking resemblances to our corpus data. Consider examples (15–17).

(15) Myn beminde Soon Carel nae onse groetenissen soo weet dat wy / Gode my beloved son Carel after our greetings so know that we God 
zy lof / noch alle welvarende zijn / verhopende ’t selve mede van u be praise still all healthy are hoping the same also of you
‘My beloved son Carel, after our greetings, you should know that we are all still in good health, God be praised, hoping the same for you, too’

(16) Seer eerwaerdighe Vader ende Moeder / Salut. Verhopende u.e. gesontheydt very honourable father and mother salut hoping your health 
en welvaert als eyghen / Godt hebbe lof and well-being as own God have praise
‘Very honourable father and mother, salut. Hoping that you are healthy and well, as I am healthy and well. May God be praised’

(17) Monsieur en seer eerwaerdige Vader ende Moeder / Salut. Ick vare wel Sir and very honourable father and mother salut I fare well 
Godt lof / verhope ’t selve van u.l. te verstaen / Dit God praise hope the same of you to understand this
tegenwoordighe sal kortelijck dienen / om u.l. t’ adviseren present will briefly serve to you to inform
‘Sir and very honourable father and mother, salut. I am doing well, praise God. Hope to hear the same from you. This one will serve to briefly inform you’

While there are striking similarities with the letters in our corpora, there are also a number of differences. Note, for instance, the use of welvarend “healthy” (15), welvaert “well-being” (16) and vare wel “am doing well” (17). Contrary to our letters, de la Chambre seems to prefer these lexemes rather than the fixed phrase kloek en gezond “strong and healthy” which predominates in our corpus. In (16), there is also the expression als eyghen, not encountered in our corpus, where similar phrasings as those preceding als eyghen could very well have triggered the use of this brief expression. Note, furthermore, the use of salut and monsieur, indicating the influence of French in the work of de la Chambre. As with Smyters, we find letters containing formulae similar to our corpus data, but as with Smyters, there are also remarkable discrepancies. One last and remarkable example: the highly frequent formula een vriendelijke groetenisse “a friendly greeting” is not found in de la Chambre’s writing manual.

The so-called materieboekjes “books of matter” were small books of some twenty pages, containing, amongst other things, examples of forms of address and a handful of example letters (de Planque 1926: 143–150). They were printed throughout the
Early Modern period, from the late sixteenth well into the eighteenth century. We will discuss two seventeenth-century editions. The first, printed in Utrecht in 1614, is the oldest surviving edition. The second edition was printed in Amsterdam in 1686. While it actually postdates the seventeenth-century letters in our corpus from the 1660s–1670s, it may still give insight into the current style of letter writing in Amsterdam, the city to which many of our letters are linked.

Figure 5. Title page of *Materi-Boecxken* 1614.

The 1614 edition from Utrecht is entitled *Materi-boecxken, oft voorschriften/seer bequaem voor die joncheyt/om wel te leeren lesen/schryven/ende een aenporringhe tot alle deuchden* “book of matter, or prescriptions, very suitable for the youth to learn to read and write, and an incentive to all virtues”. It is primarily a collection of moral lessons, addressing issues such as *Overdaet van spijs ende dranck salmen schouwen* “one should avoid excess of food and drink”, and *Van de vreese Godts* “About the fear of God”. These lessons typically take up one page, and as the book’s subtitle indicates, they may be used for both reading and writing instruction. The majority of these lessons are in plain, instructive prose, but one of them is rendered as a letter. The letter is addressed to the writer’s parents, who are requested to buy *sommige Catholycke Boecxkens* “some Catholic books”. Its beginning runs as follows:

(18) *Eersame seer beminde Vader ende Moeder ick ontbiede u hondert thausend goede nachten. En ick laet u weten dat ick noch cloec thousand good nights and I let you know that I still strong ende gesont ben (God heb +subj praze) so ick verhope dat ghy mede and healthy am God have+subj praise as I wish that you also*
zyt, den waeromme van desen is, dat
are the reason of this is that
‘Honourable, very beloved father and mother. I commend a hundred thousand good nights to you. And I let you know that I am still strong and healthy (may God be praised), as I wish that you also are. The reason of this [writing, letter] is that’

Again, there are striking similarities with the formulae found in our letters, where we also find the adjectives eersame “honourable” and beminde “beloved” (Chapter 3, Sections 6.1.1 and 6.1.3), the text-structural formula ick laet u weten dat “I let you know that” (Chapter 3, Section 6.2.1), the health formula cloec ende gesont “strong and healthy” (Chapter 3, Section 7.1), and the subordinate health formula so ick verhope dat ghy mede syt “as I wish that you also are, which I wish you are as well” (Chapter 3, Section 7.1). At the same time, there are unfamiliar phrasings such as ick ontbiede u “I commend to you” and den waeromme van desen is “the reason for this is”. These appear to be formulaic, but we have not encountered them in our letter corpora. Finally, the huge number of wishes, viz. a hundred thousand, does appear in actual usage, but not in the opening of letters as in (18), but in closings (Chapter 3, Section 6.1.4).

The 1614 Materi-boecksen also contains many examples of forms of address and of opening and closing formulae. None of these are relevant to the letters in our corpus. The examples explain how to write to noblemen such as dukes and counts, and to clergymen such as the pope and cardinals. Then the Materi-boecksen lists proper names and names of countries and animals etc., all well-known topics in elementary school books. The final part comprises examples of contracts and receipts, and four more example letters, viz. two letters addressed to the writer’s parents, and two letters addressed to the seer beminde soon “very beloved son”. In all the cases, the conclusion to be drawn is the same: there are similarities as well as differences. We will refrain from discussing the details, as the examples are similar to (18). Moreover, the example letters in the Materi-boecksen are similar to those in the more extensive letter-writing manual by Jacobi, which will be discussed below.

Before turning to Jacobi’s manual, we present one more example, taken from the second ‘book of matter’ under discussion, viz. Matery-boecksken, printed in Amsterdam in 1686. As can be gathered from the title, it is quite similar to the earlier edition from Utrecht: Matery-boecksken, ofte voor-schriften/seer bequaem voor de jonckheydt/om wel te leeren leesen/schryven/ende een opweckinge tot deughden “book of matter, or pre-scriptions, very suitable for the youth to learn to read and write, and an encourage-ment to virtues”. By way of illustration, we give here the beginning and ending of one of the example letters, which are all identical to those in the earlier edition of 1614, apart from minor spelling details and the date at the end of a letter.

(19) **Laus Deo semper.**
Praise to-God always
Lieve seer beminde Vader ende Moeder / ick recommandere my
dear very beloved father and mother I recommend me
Dear, very beloved father and mother. I humbly recommend myself into your good grace. And I let you know that I am still strong and healthy (praise God), as I hope that you are as well. The reason of this [letter] is that I

Written in Haarlem, 16 February 1685, by me N.N., your beloved son, whatever I am capable of”

Many resemblances between this sample and actual usage stand out. As with Smyters, de la Chambre and the 1614 ‘book of matter’, however, we also notice divergences, such as ick recommandere my ootmoedelijck in u goede gratie “I humbly recommend myself into your good grace” and Den waeromme van desen/dat ick “The why/reason for this, that I”. Also note the final formula wat ick vermach “whatever I am capable of”, which was disappearing from usage (see Chapter 3, Section 6.1.4), but which is nevertheless prescribed in this example.

De Planque (1926: 143, 149–150) notes many similarities between the materie-boekjes and the work of Jacobi, to the extent that the anonymously published materie-boekjes and Jacobi’s manual were clearly related in some way. Since Jacobi’s work was more elaborate than the materieboekjes, and since Jacobi’s letter-writing manual was reprinted much more often (van den Berg 1978: 4) than those of Smyters and de la Chambre, we will now concentrate on Jacobi, and discuss in more detail than before the relationship between theory and practice.

6.3 Jacobi and his successors

Jacobi’s Ghemyne zeyndt-brieven “common sending letters” was arguably the most important letter-writing manual of the Early Modern period; in any event, it had the most reprints. It was published throughout the Netherlands in a period spanning more than two centuries. The first edition was published at least as early as 1597, in Amsterdam, the latest edition dates from Maastricht (Limburg), 1803. For the present

8. De Planque (1926: 134–143) assumes that the 1597 Protestant edition is not the oldest one, but that it was founded on an earlier Catholic edition.
Firstly, we noticed the great similarity of these four editions. Apart from minor changes, the macro-structure of the book had not altered in the course of time. This also applies to the micro-structure, which means that most example letters from the edition of 1645 verbatim appear in the edition of 1792 as well.

Figure 6. Title page of Jacobi’s *Gemeyne zeyndt-brieven* 1645.
Having established these great similarities, we focused on Jacobi’s 1645 edition, which was printed in Holland, the origin or destination of most of our letters, and just before the period from which our seventeenth-century letters stem (1660s/1670s).

In Table 8, we give a representative selection of fourteen epistolary formulae found in the instruction and in the seventy example letters in prose in Jacobi (1645), indicating their frequencies in the right-hand column. We distinguish between formulae that are also found in actual usage in the letter corpora used for Chapter 3, and formulae that we have not found there.

### Table 8. Epistolary formulae in Jacobi (1645) with their frequencies.

<table>
<thead>
<tr>
<th>Formula</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Found in actual use</strong></td>
<td></td>
</tr>
<tr>
<td>1 laus deo semper</td>
<td>4</td>
</tr>
<tr>
<td>‘praise God forever’</td>
<td></td>
</tr>
<tr>
<td>2 een vriendelijke groetenisse zij geschreven</td>
<td>2</td>
</tr>
<tr>
<td>‘a friendly greeting be written’</td>
<td></td>
</tr>
<tr>
<td>3 na vriendelijke groetenisse</td>
<td>1</td>
</tr>
<tr>
<td>‘after friendly greeting’</td>
<td></td>
</tr>
<tr>
<td>4 kloek en gezond (3), kloek ende sterck (1), wel te passe (1), in goeden doene (1)</td>
<td>6</td>
</tr>
<tr>
<td>‘(strong and) healthy’</td>
<td></td>
</tr>
<tr>
<td>5 God heb lof</td>
<td>6</td>
</tr>
<tr>
<td>‘Praise the Lord’</td>
<td></td>
</tr>
<tr>
<td>6 zo ik verhope dat gij zijt/van u te horen/vinden</td>
<td>5</td>
</tr>
<tr>
<td>‘as I hope you are/to hear from you/to find about you’</td>
<td></td>
</tr>
<tr>
<td>7 dat ik u brief ontvangen heb en daaruit verstaan</td>
<td>3</td>
</tr>
<tr>
<td>‘that I have received your letter and understood from it’</td>
<td></td>
</tr>
<tr>
<td>8 laat u weten dat</td>
<td>10</td>
</tr>
<tr>
<td>‘let you know that’</td>
<td></td>
</tr>
<tr>
<td>9 zijt/blijft hiermede den here bevolen</td>
<td>42</td>
</tr>
<tr>
<td>‘lit. be/stay with this the Lord commended’</td>
<td></td>
</tr>
<tr>
<td><strong>Not found in actual usage</strong></td>
<td></td>
</tr>
<tr>
<td>10 ick zende/ontbiede u duizend/een miljoen/veel goede nachten</td>
<td>11</td>
</tr>
<tr>
<td>‘I send/(re)commend you a thousand/one million good nights’</td>
<td></td>
</tr>
<tr>
<td>11 ik recommandere/gebiede mij t’uwaarts/in u gratie</td>
<td>15</td>
</tr>
<tr>
<td>‘I (re)commend myself to you/in your grace’</td>
<td></td>
</tr>
<tr>
<td>12 presenteer mijn gedenstigheydt t’uwaarts</td>
<td>1</td>
</tr>
<tr>
<td>‘present my helpfulness to you’</td>
<td></td>
</tr>
<tr>
<td>13 oorzake mijns schrijvens is</td>
<td>7</td>
</tr>
<tr>
<td>‘reason of my writing is’</td>
<td></td>
</tr>
<tr>
<td>14 den waaromme van dezen is</td>
<td>8</td>
</tr>
<tr>
<td>‘the reason of this is’</td>
<td></td>
</tr>
</tbody>
</table>
As is clear from Table 8, the manual contains a range of formulae also found in actual usage. Most of these formulae appear only a few times in Jacobi, whereas they occur much more frequently in our letter corpora. It should be noted, for instance, that the seventeenth-century subcorpus used for Chapter 3 comprised 81 letters, while Jacobi (1645) contained 70 letters and a short introduction with instructions. But where actual usage provided us with 39 tokens of formula 2 (een vriendelijke groetenisse zij geschreven), with 43 tokens of formula 4 (kloek en gezond), and with 64 tokens of formula 8 (laat u weten dat), the frequencies in Jacobi are 2, 3 and 10, respectively. Only the commendation formula (formula 9 in Table 8), with its 42 tokens in Jacobi's manual, is about equally frequent in actual usage, where we found 51 tokens. Obviously, as a manual, Jacobi's booklet will have been aimed at offering a wide variety of linguistic choices, providing as many resources as possible to letter writers. It would certainly be wrong to expect too great a uniformity in Jacobi's sample letters. Still, these striking differences in frequency between theory and practice leave unexplained why some formulae are so frequent in actual usage, while they are only one of the many options Jacobi offered.

Interestingly, Jacobi also provides sample letters with formulaic health passages. Note that formulae 4, 5 and 6 in Table 8 occur five or six times in his manual, and in fact they occur five times in the same sample letters, and in the same order. Clearly, the formulaic discourse on health attested in actual usage also appears in Jacobi's manual, and as in the letter corpora, these strings of health formulae and subordinate health formulae occur in the introductions to letters. But whereas actual usage provides us with three more subordinate health formulae (see Table 11 in Chapter 3, Section 7.1.3), repeated here as (20, 21, 22), these do not occur in Jacobi (1645).

(20) het welcke mijn van harten seer lief om te hooren is
‘which I would very much love to hear’

(21) waer het Anders het waer ons van herten leet
‘if it were different, we would be very sorry’

(22) dat wedt godt almachtich die en kender van alle harten is
‘The almighty God, who knows all the hearts, knows this’

Many more such discrepancies exist. Formula 1 in Table 8, laus deo semper, occurs in actual usage, but its Dutch alternative loot godt boven al “praise God above all” (example (8a) in Chapter 3, Section 6.1.2) is much more frequent, although it is not present in Jacobi's manual. Similarly, a formula such as nijet mer op dijt pas dan “nothing more for now but” (example (18a) in Chapter 3, Section 6.2.3), creating the transition to the closing of a letter, is fairly frequent in actual usage, but absent in Jacobi (1645).

Over a third of Jacobi's sample letters, on the other hand, begin with formula 10, 11 or 12 from Table 8, but not a single letter in our letter corpora begins with any of these formulae. The expression duizend goede nachten “thousand good nights” does occur, but as a greeting or closing formula in the endings of letters (see Chapter 3, Section 6.1.4). Likewise, Jacobi marks the transition from the opening to the middle
part fifteen times with either formula 13 or 14. The total corpus used for the present monograph contains exactly one example of formula 13, and none of formula 14.

As explained above, Jacobi obviously offered a variety of linguistic choices. Still, he is quite consistent in two respects. The first is that the majority of letters end with a commendation formula such as formula 9 in Table 8. This is no different in actual usage. The second is that almost all letters in Jacobi are dated at the end, close to the signature. In Chapter 3, Section 6.1.2 (Table 4), we showed that most of the seventeenth-century letters used for that chapter (68.3%) were dated at the beginning. Towards the end of the eighteenth-century, however, a small majority (56.0%) were dated at the end. Even if this change was influenced by Jacobi, it does not contribute greatly to the strength of Jacobi’s influence: 120 years later, and almost half of the letter writers still dated their letters at the beginning.

Finally, it is interesting to note that whereas actual writing practices evolved, creating sheer differences between the seventeenth and the eighteenth centuries, Jacobi’s manual was printed over and over again, remaining largely the same. We compared the editions of 1656, 1728 and 1792 with the 1645 edition discussed above. Strikingly, whereas both printing and handwriting evolved in this period, the later editions of Jacobi were all still printed in the same sixteenth-century civilité which was used for its earliest editions. The contents of the book remained equally stable throughout the entire period, although the dates and the names below the example letters were sometimes modernised. The 1792 edition from Venlo, for instance, contains letters supposedly written in Antwerp and Bruges in the late eighteenth century, which were written in Amsterdam and Haarlem in the early seventeenth century according to the 1645 edition from Hoorn. The spelling was also modernised: in Jacobi (1645: A5r), a letter is closed with the formula Wat ick vermach “whatever I can”; in the 1792 edition we find Wat ik vermag (p. 11), with ⟨k⟩ instead of ⟨ck⟩, and ⟨g⟩ instead of ⟨ch⟩. Much more salient than these common orthographical innovations, however, is the fact that the formula itself is still there. Although this well-known formula was probably already on the decline in the seventeenth century (as explained in Chapter 3, Section 6.1.4), it was apparently still recommended to letter writers in Venlo, in the year 1792. Similarly, Jacobi’s Laus Deo semper (formula 1 in Table 8; 1645: B1v) and Een vriendelijke groetenisse geschreven (formula 2 in Table 8; 1645: B4v), both of which vanished from usage after the seventeenth century, survived in the 1792 edition (p. 23 and 29 respectively). The same applies to kloek en gezond (formula 4 in Table 8) which disappeared from usage and was partly replaced by fris en gezond “fresh and healthy” in the eighteenth century. It does, however, occur in both Jacobi (1728) and Jacobi (1792), while fris en gezond does not.

The increasing divergence between Jacobi’s manual and actual usage did not remain unwitnessed. Two letter-writing manuals were published, aimed at replacing Jacobi, to a large extent by modernising it. In the introduction to his Send-brieven (1740), Koolenkamp explicitly stated that he had adopted much of Jacobi’s manual and adapted its contents to de hedendaagse stijl “the current style”, because many of Jacobi’s phrasings were niet gebruikelijk “not common” anymore (A1v). By this, Koolenkamp
clearly must have meant that he had removed most of the epistolary formulae occurring in Jacobi’s manual. In the 41 example letters that Koolenkamp offered, following letter-writing instructions taken almost verbatim from Jacobi, he used very little formulaic language. Only occasionally do we find formulae that also occur in our corpora. One example is *dese sal dienen* “this will serve” (B2v; see Chapter 3, Section 6.1.3), variants of which formula occur 17 times in the 100 eighteenth-century letters used for Chapter 3. Another example is *Eindigende blijve na aanbiedinge van mijn dienst* “lit. ending, (I) remain after offering my service” (B4v), which appears to be a typically eighteenth-century closing formula (see Chapter 3, Section 6.2.3). As with Jacobi, there are two obvious inferences to be derived from Koolenkamp’s manual. Firstly, that letters should be concluded with a formula containing *Blijve* “(I) remain”, as in the previous example. Koolenkamp ended most of the letters in this way, thus closely resembling actual usage. Secondly, that letters should be dated at the end. As we have seen previously, this was also common practice in the various editions of Jacobi’s book, and eighteenth-century usage showed a small preference for dating at the end.

Much of what has been said in connection with Koolenkamp also applies to Hakvoord. Where Koolenkamp wrote his book for *de Jeugt* “the young”, Hakvoord intended his *Gemeene zend-brieven* “Common sending letters” to be used in *de Duitse scholen* “in the Dutch schools” (title page), referring in fact to the lowest level of primary education. Contrary to Koolenkamp, Hakvoord’s manual was printed many times in the eighteenth century (van den Berg 1978: 4–5). We examined two editions, one from ca. 1750 and one from 1794, which proved to be almost identical. The book contains only 21 example letters following five pages with instructions. It is noteworthy that Hakvoord commented on the place of the date: it should be put at the end of the letter, not at the beginning, *hoewel ’t van eenige Kooplieden, nog somtijds word gedaan* “although this is still sometimes done by some merchants” ([1750]: 7). The eighteenth-century letters in our corpus are far from being all written by merchants, but still 44% of the letters examined for Chapter 3 are dated at the beginning. Similar discussions are found in English manuals of the seventeenth and eighteenth centuries. Nevala (2004b: 39) notes that these manuals tend to prescribe that the date should be placed at the bottom of the letter, but at the top when writing to a superior. While this appears to be at odds with the Dutch situation, it should be pointed out that the opposite was also argued for (Nevala 2004b: 39). Another similarity with both Koolenkamp and actual usage is the repeated occurrence of closing formulae containing elements of the commendation formula and/or *blijve* “remain”. Note also that all example letters repeat the addressing of the recipient, which was shown to be a characteristic of eighteenth-century usage (Chapter 3, Section 6.1.4).

Both successors to Jacobi, viz. Koolenkamp and Hakvoord, showed some resemblances with eighteenth-century usage. In so far as epistolary formulae were largely absent from their works, there was also a clear difference from actual usage, albeit that the use of formulae sharply diminished over time (see Chapter 4). In fact, it could be argued that the most obvious resemblance between these eighteenth-century manuals and actual usage was that neither contains very many formulae, or in any case they
contain far fewer formulae than seventeenth-century manuals and letters. Evidently, this implies that any direct influence of the theory of letter writing on its practice cannot be demonstrated. At most, we can observe that both in theory and in practice the language of letters tended to become less formulaic, one final caveat being that Jacobi’s much more formulaic manual remained in print throughout the period, and was even reprinted in the early nineteenth century.

These writing manuals with a fairly modest or bourgeois target audience and/or an orientation to schools were all written by schoolteachers or explicitly intended for use in schools. Whereas it may be assumed that schoolteachers such as Smyters and de la Chambre used their own manuals when teaching, there is scant evidence that this was done on a larger scale. De Booy, in her studies of primary education in the province of Utrecht, notes that even the small (and therefore inexpensive) materieboekjes “books of matter” were probably not widely used, viz. only by the (minority of) pupils who were taught writing skills (1977: 57). The materieboekjes were not mentioned in school regulations nor on schoolteachers’ receipts (De Booy 1977: 276). De Booy assumes that the manuals by Jacobi, de la Chambre, Koolenkamp and Hakvoord were mainly used by children who would become professional scribes or secretaries (De Booy 1977: 276). These manuals were not commonly owned by pauper schools, and their use may have been largely restricted to private schools (De Booy 1980: 48). Particular remarks in their prefaces, mentioning the possibility of translating the example letters into Latin and French, indicate that Jacobi and the materieboekjes were also intended to be used in secondary education (de Planque 1926: 135, 144; Jacobi 1645: a1v; materieboekjes 1614, 1686). In short, there is little evidence that any of the manuals under discussion were widely used in elementary schools, while it is certain that many other books such as basic reading manuals and catechisms were used on a very large scale, that is, in every school (cf. de Planque 1926; de Booy 1977, 1980; Roosenboom 1997).

Summing up, we have established resemblances and differences between theory and practice. What does this imply for the practice of letter writers and the question of how they acquired their formulaic language? In our view, the differences outweigh the resemblances to the extent that no direct influence of letter-writing manuals on actual practice should be assumed (cf. also van der Wal & Rutten 2013b). Although the manuals present letter-writing conventions, it is more likely that pupils at school and youngsters at home learnt letter-writing conventions and formulaic language and further developed their writing skills by participation in actual writing practices, that is by reading and copying letters and hearing them read aloud. In the same vein, Austin (1973a: 12) argued that the differences between actual language use in English letters from the late eighteenth-century and contemporary theory were surprisingly great.

9. Schoolteachers’ receipts mention the books schoolteachers bought for poor children, for which they would be reimbursed by the parish. The two most popular books by far were the catechism and elementary alphabet books (de Booy 1977: 271). The more advanced manuals were probably outside the scope of poor relief pupils. Similarly, Roosenboom (1997: 228) notes that neither materieboekjes nor writing manuals were mentioned on a 1767 reading list for the Den Bosch area in the province of Brabant.
Epistolary conventions were probably more strongly steered and transmitted to others through writers’ participation in the practice of letter-writing, that is, by receiving and reading letters, and by drafting replies (cf. Austin 1973a: 13). Similar claims have been made for English letter-writing in the long eighteenth century (Whyman 2009: 28–45; Brant 2006: 9–10), and for German and Finnish letters from the nineteenth century (Elspaß 2005: 194–195; Laitinen & Nordlund 2012), as well as for Dutch elite correspondence from around 1800 (Ruberg 2005). Actual usage is always more varied than reconstruction on the basis of letter-writing manuals would allow. Nevertheless, the manuals bear witness to some of the conventions that were in use, and to their authors’ need to lay down in books the conventions with which many people were already familiar. As such, they may still have functioned as books of reference.

7. Conclusions

Our exploration of various aspects of the writing process leads us to the following conclusions. Firstly, the frequency differences between the autograph subcorpus and the non-autograph subcorpus prove that social and professional writers made even more use of formulae than autograph writers. Writing in the name of someone else appears to trigger the use of formulaic language. Again, this gives more insight into the social practice of everyday letter-writing.

Secondly, we may conclude that our assumption of epistolary formulae as a safe haven for writers suffering from communication problems has been proved correct. As we have seen, these problems typically originate from two problematic communicative situations in which writers can be involved. On the one hand, writers, who are also the sender, may have little writing experience, and thus suffer from form-oriented problems. They may know what to say, but they do not know how. With the use of formulae, they fall back on standard forms, and thus on standard contents. On the other hand, social or professional writers, who are not identical to the sender, have to formulate contents which are not their own. With the use of formulae, they fall back on standard contents, and thus on standard forms. In both cases, the ultimate result is a highly formulaic letter.

Finally, we also established resemblances to normative publications on letter writing, although we doubted any direct influence of these manuals on actual language use. In our view, formulaic language did not abound because normative publications prescribed the use of formulae, but rather because formulae fulfilled clear functions.
CHAPTER 6

Forms of address

1. Introduction

Forms of address probably belong to the most salient characteristics of historical letter writing. They are certainly among the most well-researched topics of historical letter writing. As in many other languages, the Dutch letters in our corpus display an impressive variability with regard to forms of address, both synchronically and diachronically. In this chapter, we will show that the variation found can largely be explained by sociolinguistic factors.

In Section 2, we will give a brief overview of the history of forms of address in Dutch, thereby introducing the various nominal and pronominal forms attested for the seventeenth and eighteenth centuries. Forms of address in historical discourse have mainly been studied from a pragmatic perspective, with a strong focus on politeness theory. Therefore, we will seek to explain the variation in the use of forms of address in our corpus from the view of politeness which is the topic of Section 3. We will approach the matter from various perspectives, drawing the conclusion that traditional politeness theory on the one hand helps to explain some of the variation found, but on the other hand leaves important issues without any explanation. In Section 3.5, we will therefore tie in with Watts’ (2003) redefinition of politeness as politic behaviour, refraining from initial judgements on the either negative or positive character of specific variants. This gives way to an interpretation of the different forms as variants in the sociolinguistic sense, which implies that their use is open to sociolinguistic research. In Section 4, we will show that a sociolinguistic approach drawing on social class, gender and time as the main external variables can explain much of the variation found. Nevertheless, some patterns remain difficult to understand, and in Section 5 we will argue that politeness strategies do seem to influence the distribution of particular variants, but only in very specific contexts.

In Section 4, where we approach the distribution of forms of address from a sociolinguistic perspective, we examine the seventeenth-century (autograph) subcorpus and the eighteenth-century subcorpus, the basic numbers of which were presented in Chapter 1 (Section 6). For the detailed composition of these corpora we refer to Chapter 4 (Section 3.1). These corpora were used in Chapter 4 and will be used in all following case studies. In Section 3, where we approach the variation drawing on politeness theory, we will occasionally use smaller subcorpora, which will be introduced in the appropriate sections.
2. Forms of address in the history of Dutch

Since in the present chapter we focus on pronominal forms of address and on nominal ones commonly behaving as pronouns, we first introduce the different forms and provide some historical-linguistic background in this section.

The history of second person pronouns is as complicated in Dutch as it is in many other languages. In most Dutch dialects, the historical second person singular pronoun *du* “you”, similar to German *du* and English *thou*, had given way to *gij* “you” by the seventeenth century (Berteloot 2003). The form *gij* was originally a second person plural pronoun, which entered the singular paradigm, as happened in many languages, for instance English *you*, French *vous*, Czech *vy*, Finnish *te*, Turkish *Siz* and Swedish *ni* (cf. Jucker & Taavitsainen 2003: 4). In most published texts from the seventeenth and eighteenth centuries, *gij* and variants such as *gy*, *ghy*, *ghij* are the common if not the only second person pronouns, in both the singular and the plural (cf. van der Horst 2008: 1093).

But the system was heavily in flux, and many more variants were in use. In the singular, *jij* was used, in particular in texts from the northern parts of the language area, including Holland, to which many of our letters are related. This form is already found in early-seventeenth-century texts, and is often assumed to have been the regular variant in the spoken language of the Northern Netherlands (van der Horst 2008: 1093). In written language, however, *gij* dominated well into the nineteenth century, including in texts from the north, while the use of *jij* in the seventeenth and eighteenth centuries was largely restricted to informal settings in plays (Vermaas 2002: 42–52). In present-day standard Dutch, *jij* is the so-called T-form, whereas *u* is the V-form. The oldest use of *u* as a subject also dates back to the seventeenth century (van der Horst 2008: 1094), but it hardly entered the written language of the seventeenth and eighteenth centuries, and was still outnumbered by *gij* in the nineteenth century (van der Horst 2008: 1374, 1652).

While *jij* and *u* alternated with *gij* in the singular, variants of plural *gij* also came into use. Composite forms such as *gijlieden* “you people”, already found in Middle Dutch, were also in use in the sixteenth, seventeenth and eighteenth centuries (van der Horst 2008: 814, 1096, 1375–1376). In the eighteenth century in particular, composite forms such as *jijlui*, *jelui* and *joului* were quite frequent (van der Horst 2008: 1375), consisting of *jij* or one of its variants, nominative or oblique, and the suffixoid *lui*, a variant of *lieden* “people”. In present-day standard Dutch, one such composite form, viz. *jullie* “you”, may be characterised as the plural T-form.

So far, we have mentioned the full nominative forms of second person pronouns commonly found in written (Early) Modern Dutch, abstracting away from not only the wide range of present-day dialect and substandard forms (cf. de Vogelaer 2008), but also from both reduced forms and oblique forms. We will return to this below, but

1. See e.g. the papers in Taavitsainen & Jucker (2003) and in Culpeper & Kádár (2010).
first we should discuss nominal forms of address, which are pervasive in Early Modern Dutch, and certainly in letters.

Apart from the pronominal forms of address, nominal forms occurred consisting of a possessive or oblique pronoun and a respectful title. Again, there are many similarities with other languages (Jucker & Taavitsainen 2003: 4). In Spanish, for instance, *Vuestra Merced* “Your Honour” grammaticalised into the V-form *Usted*, as did *la Vosstra Signoria* into *Lei* in Italian. In historical Dutch, combinations of a possessive or oblique pronoun and a noun occur in many instances, the two dominant variants being *ul* and *ue*. The first form, *ul*, is usually considered an abbreviation of *uwe liefde* or *u liefde* “your love, your kindness”, occurring mainly in fifteenth-, sixteenth- and seventeenth-century Dutch (*MNW* s.v. *liefde*, *WNT* s.v. *liefde*). Alternatively, *ul* could refer to *u lieden* “you people” (*WNT* s.v. *ulieden*). The first could function as a singular and a plural, the second obviously only as a plural. The other form, *ue*, is commonly interpreted as *u edele* or *uwe edelheid* “your honour”, and appears to have been in use from the sixteenth well into the nineteenth century (*WNT* s.v. *ue*). Both *ul* and *ue* were used in all cases, and also as a possessive.

Both *ul* and *ue* are usually considered as prototypically epistolary forms, and their impressive frequency in our letter corpora certainly seems to confirm this idea. At the same time, however, *ul* and *ue* are often assumed to have influenced the rise of *u* as a new second person (subject) pronoun, or even to have been the source of this new form. In other words, they may have been in use as general forms of address in both spoken and written language. For reasons of convenience, we will nevertheless refer to *ul* and *ue* as epistolary forms of address, subsuming all other forms discussed under the label of pronominal forms of address.

On the basis of the existing literature,² the following table (Table 1) can be drawn of the forms of address that are frequently mentioned in studies of seventeenth- and eighteenth-century Dutch (in normalised spelling). We distinguish between nominative, oblique and possessive forms. Oblique and possessive forms are identical to nominative forms in the case of *ul* “your love” and *ue* “your honour”, and similar, i.e. with identical initial consonants, in the case of *d*-forms and *j*-forms. In the case of *gij*, however, the oblique and possessive are *u(w)*, of which the oblique is often considered another possible source of the above-mentioned rise of nominative *u*. Furthermore, the table comprises both the full and the reduced forms, but it should be noted that there were probably conditions on the use of either, as there are nowadays. To give an example, the full form *jij* [*jei*] counts as the present-day standard T-form, but in spoken language its reduced variant *je* [*jo*] is much more common in unstressed positions, where *jij* is never strictly impossible, but nevertheless may sound somewhat out of place (*ANS* 5.2.7). We also give the present-day standard variants as a point of reference.

². See van den Toorn (1977), Berteloot (2003), van der Horst (2008), Aalberse (2009) and the references there.
Table 1. Forms of address in seventeenth- and eighteenth-century Dutch, and in the present-day standard variety.

<table>
<thead>
<tr>
<th></th>
<th>Nominative</th>
<th>Oblique</th>
<th>Possessive</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>17th and 18th centuries</strong></td>
<td><strong>Sing Pronominal</strong></td>
<td>du, dou</td>
<td>di</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gij, ge</td>
<td>u</td>
</tr>
<tr>
<td></td>
<td></td>
<td>jij, je</td>
<td>jou, je</td>
</tr>
<tr>
<td></td>
<td></td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td></td>
<td><strong>Epistolary</strong></td>
<td>ul</td>
<td>ul</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ue</td>
<td>ue</td>
</tr>
<tr>
<td><strong>Plur Pronominal</strong></td>
<td>gij, ge</td>
<td>u</td>
<td>uw</td>
</tr>
<tr>
<td></td>
<td></td>
<td>gjilieden</td>
<td>ulieden</td>
</tr>
<tr>
<td></td>
<td></td>
<td>joului, jelui</td>
<td>joului, jelui</td>
</tr>
<tr>
<td></td>
<td></td>
<td>jullie</td>
<td>jullie</td>
</tr>
<tr>
<td></td>
<td><strong>Epistolary</strong></td>
<td>ul</td>
<td>ul</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ue</td>
<td>ue</td>
</tr>
<tr>
<td><strong>Present-day standard</strong></td>
<td><strong>Sing</strong></td>
<td>jij, je</td>
<td>jou, je</td>
</tr>
<tr>
<td></td>
<td></td>
<td>u</td>
<td>u</td>
</tr>
<tr>
<td></td>
<td><strong>Plur</strong></td>
<td>jullie</td>
<td>jullie</td>
</tr>
</tbody>
</table>

By the seventeenth century, the historical d-forms had largely vanished from the regions of Holland and Zeeland, and we therefore did not expect to find many examples of these in our letters (and indeed our findings bear this out). Because of the nature of our sources, viz. private letters, many of which were sent to close family members, one might expect j-forms such as jij, je, jou and jullie, but these turn out to be quite scarce in our letters. In fact, there are very clear patterns of usage, despite the wide range of variation displayed in Table 1. In the seventeenth century, four variants, viz. gij, ul, ue and u (oblique and possessive), dominate. In the eighteenth century, three of these, gij, ue and u (oblique and possessive) continue to prevail. All other forms are marginal.

Before dealing with the forms of address found in our letter corpora, we have to make a comment on the variable of region. One of the strongest predictors of pronominal usage in present-day Dutch is region (cf. de Vogelaer 2008), and one might expect region also to have been an important variable in the past. Although this may apply to the spoken language of the past, it does not pertain to the written language of our letters. The pronominal forms gij and u and the nominal forms ul and ue are used throughout the regions represented in our corpus, without significant differences, for instance, between letters from Holland and letters from Zeeland. At one point, however, regional differences do seem to be important, though not differences between Dutch dialect areas, but between letter writers from the Netherlands and writers from
the “abroad” category in our corpus, viz. those who were born and/or raised in Dutch-speaking families living abroad (see Section 3.3.2). The virtually absent effect of region on the use of forms of address reminds us once more of the fact that we are dealing with a written variety of Dutch with many supraregional characteristics.

3. Politeness as a variable

Most historical-linguistic research on forms of address focuses in one way or another on politeness (e.g. Taavitsainen & Jucker 2003; Nevala 2004b), following to a greater or lesser extent the well-known model of Brown & Levinson (1978, 1987). In this section, we will discuss to what extent we believe politeness to be a fruitful approach for the use of forms of address in our letter corpora by concentrating on cases where the influence of politeness strategies would be expected.

A thorough overview of politeness research in a historical-sociolinguistic context is given by Nevala (2004b: 54–80). Maintaining the distinction between strategies of positive and negative politeness, she describes positive politeness in terms of intimacy and solidarity (2004b: 59–60). The corresponding strategy with regard to forms of address involves the singular non-honorific pronoun. In our letters, the \textit{j}-forms would be positive polite forms, as would perhaps \textit{gij}, although its exact value in terms of politeness is obscured by the fact that it may not have been widely used in the spoken language of the northern/north-western parts of the Netherlands in the seventeenth and eighteenth centuries. Nevala (2004b: 60–61) discusses negative politeness in terms of social distance, indirectness and deference, and the corresponding politeness strategy would involve honorifics. In our corpora, the nominal forms of address \textit{ul} “your love” and \textit{ue} “your honour” would be honorifics. In the present section, we will use the same distinction, considering \textit{ul} and \textit{ue} as relatively negatively polite forms of address, and all other forms as relatively positively polite.

In Section 3.1, we contrast a subset of seventeenth-century private and business letters, and in 3.2, we compare the results by focusing on forms of address in the most intimate relationships, viz. within the core family. In 3.3, we concentrate on what were probably unambiguous T-forms, viz. the \textit{j}-pronouns. We then turn to the use of epistolary forms, assuming for the moment that these are more negatively polite than pronominal forms of address (3.4). Our conclusion in many instances is that a politeness approach does yield some valuable insights, but does not provide a satisfactory explanation for the variation found. In Section 3.5, we, therefore, argue for an approach in terms of politic behaviour instead, following among others Watts (2003). In such an approach, the forms of address are primarily interpreted as conventional signs, not necessarily invoking either negatively or positively polite associations. The approach also allows for a social stratificational investigation, to which we will turn in Section 4.
3.1 Private vs business

As an initial approach, we compared the forms of address found in private letters with those in business letters. Even though our prime interest lies with private letters, our seventeenth-century corpus also contains a fair number of business letters allowing for a well-founded comparison of private and business correspondence. If any significant differences in the use of forms of address were to be expected, such differences would certainly be apparent in this comparison of private and business letters. As a reasonable assumption, business letters would show more negative politeness strategies than private letters, and hence more nominal forms of address than pronominal forms. As most business letters in our corpus were written by UMC men, we have neutralised the social variables of class and gender and focus here on seventeenth-century private and business letters from UMC men. We selected all letters written by UMC men from the seventeenth-century autograph subcorpus, both private and business, and extracted all forms of address. The results are presented in Table 2.

As becomes clear from Table 2, the main variants are _ul, ue, gij_ and _u_. The epistolary forms _ul_ and _ue_ occur in all positions, while _gij_ is restricted to the nominative and _u_ to the oblique and possessives. Historically, _u_ is the oblique form of the pronoun _gij_, which is the nominative form. Forms with _j_-anlaut as well as nominative _u_ are infrequent variants in both private and business letters. Most _j_-forms are found in private letters, which may be explained by their relative informality. The relative novelty of _u_ as a subject may have rendered the form less appropriate in written language in general.

<table>
<thead>
<tr>
<th>UMC-m, 17th century (autographs)</th>
<th>Private letters 121 letters, 86 writers</th>
<th>Business letters 41 letters, 35 writers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td><em>ul</em> (nom/obl/poss)</td>
<td>501</td>
<td>40</td>
</tr>
<tr>
<td><em>ue</em> (nom/obl/poss)</td>
<td>333</td>
<td>27</td>
</tr>
<tr>
<td><strong>Subtotal epistolary forms</strong></td>
<td>834</td>
<td>67</td>
</tr>
<tr>
<td><em>gij</em> (nom)</td>
<td>155</td>
<td>12</td>
</tr>
<tr>
<td><em>u</em> (obl/poss)</td>
<td>238</td>
<td>19</td>
</tr>
<tr>
<td><em>j</em>-form (nom/obl/poss)</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td><em>u</em> (nom)</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1,250</td>
<td>100</td>
</tr>
</tbody>
</table>

3. Here and throughout this chapter, we will use normalised spellings when discussing the forms of address, except in examples. Main orthographic variants include UL and UE with capitals and ⟨gh⟩-anlaut instead of ⟨g⟩ in the case of _gij_, and finally there is the well-known variation of ⟨i⟩, ⟨j⟩, ⟨γ⟩, ⟨ey⟩, ⟨ij⟩ et cetera.
The epistolary forms originated as indirect, nominal forms of address. If we consider these as negatively polite relative to the direct pronominal forms *gij* and *u* (obl/poss), we see a clear difference between private and business letters. Whereas the private letters contain 67% epistolary forms, the share of *ul* and *ue* rises to 89% in the business letters, where the proportion of *gij* and *u* is obviously correspondingly lower (31% vs 10%). This difference is exactly what we would expect, considering the nature of the two text types. Private letters between friends and family members tend to be more intimate than business letters exchanged by business partners. Thus there appears to be a difference between private and business letters, which could be interpreted as more positive politeness in private letters as opposed to more negative politeness in business letters. Apart from this text type difference, it appears even more striking that the epistolary forms take up no less than two-thirds of all forms of address in private letters. This would mean that the dominant politeness strategy employed in private letters between spouses, friends, siblings and the like was negative. At the same time, while the use of *gij* and *u* (obl/poss) is much more frequent in private letters than in business letters, it should still be noted that *gij* is used by no fewer than 9 out of 35 writers of business letters (26%). In other words, it is questionable to what extent *gij* should be considered only a minor variant in business correspondence. Summing up, there seem to be reasons to consider the epistolary forms as negatively polite, but there are also reasons to doubt such a categorisation. We will return to this in Section 3.5.

### 3.2 Within the core family

Focusing our attention on concrete communicative contexts might reveal more of the parameters influencing the choice of forms of address (cf. Nevala 2004b: 175–177). Being struck by the sheer amount of supposedly negatively polite forms in private letters, we therefore embarked upon a second case study in an effort to more fully understand this usage. In order to enhance comparability with the results in the previous section, we have confined ourselves to seventeenth-century autographs. We focus on the use of forms of address within the core family, that is, on the communicative situation where intimacy is expected to be more important than in any other setting, and where, accordingly, the lowest number of negatively polite forms would be expected.

Within the core family, four types of relationship provide us with a fair number of letters from the seventeenth-century autograph subcorpus, viz. parents to children, children to parents, letters exchanged between spouses, and letters between siblings. We selected all letters allocated to one of these types of relationship, and extracted all forms of address. Table 3 presents the results according to the type of relationship. As can be gathered from Table 3, the proportion of epistolary forms is lowest in letters from parents to children (48%), rises in letters between spouses (60%) and in letters from children to parents (61%), and peaks in letters between siblings (68%). The proportion of *gij* and *u* (obl/poss) is accordingly lower. Again, the proportion of *j*-forms and *u* (nom) is conspicuously low.
Letters as Loot

Table 3. Forms of address according to the relationship between the sender and the addressee.

<table>
<thead>
<tr>
<th>Core family</th>
<th>Parent &gt; child</th>
<th>Spouse &gt; spouse</th>
<th>Child &gt; parent</th>
<th>Sibling &gt; sibling</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th century (autographs)</td>
<td>24 letters</td>
<td>99 letters</td>
<td>18 letters</td>
<td>20 letters</td>
</tr>
<tr>
<td></td>
<td>13 writers</td>
<td>82 writers</td>
<td>15 writers</td>
<td>19 writers</td>
</tr>
<tr>
<td>N %</td>
<td>N %</td>
<td>N %</td>
<td>N %</td>
<td></td>
</tr>
<tr>
<td><em>ul</em> (nom/obl/poss)</td>
<td>136 38</td>
<td>718 51</td>
<td>63 39</td>
<td>148 49</td>
</tr>
<tr>
<td><em>ue</em> (nom/obl/poss)</td>
<td>36 10</td>
<td>125 9</td>
<td>35 22</td>
<td>60 20</td>
</tr>
<tr>
<td>Subtotal epistol. forms</td>
<td>172 48</td>
<td>843 60</td>
<td>98 61</td>
<td>208 68</td>
</tr>
<tr>
<td><em>gij</em> (nom)</td>
<td>83 23</td>
<td>261 18</td>
<td>21 13</td>
<td>47 16</td>
</tr>
<tr>
<td><em>u</em> (obl/poss)</td>
<td>96 27</td>
<td>284 20</td>
<td>31 19</td>
<td>45 15</td>
</tr>
<tr>
<td><em>j</em>-form (nom/obl/poss)</td>
<td>9 3</td>
<td>27 2</td>
<td>3 2</td>
<td>4 1</td>
</tr>
<tr>
<td><em>u</em> (nom)</td>
<td>2 1</td>
<td>1 0</td>
<td>9 6</td>
<td>– –</td>
</tr>
<tr>
<td>Total</td>
<td>362 100</td>
<td>1,416 100</td>
<td>162 100</td>
<td>304 100</td>
</tr>
</tbody>
</table>

The numbers for siblings correspond very well to the overall numbers of private letters by UMC men (Table 2), which could be explained by the fact that most of these letters are exchanged between adult siblings, who are embarking on their own lives and starting their own families. As a result, relative intimacy gives way to relative distance, and their verbal behaviour exhibits conformity with this. In the other three types of relationship, the proportion of epistolary forms is consistently lower than the share of 67% in private letters by UMC men (Table 2). In the case of letters from parents to children, the number even drops below 50%, which could be explained by the fact that this is the one relationship where unequal power relations may be uncontroversially assumed to have existed (see below on spouses).

As before, the effect of politeness strategies can be witnessed in the results, as can seemingly divergent patterns also. If we assume that the relative power of parents over children reduces the use of negatively polite epistolary forms to less than 50%, we nevertheless have to observe that parents writing to their children still produce epistolary forms in 48% of all cases. Their verbal behaviour does not differ considerably from that of children writing to their parents, where the proportion of *ul* is almost equal, but where *ue* rises from 10% to 22%. The imbalance of power therefore has a fairly small influence on the proportion of negatively polite forms.

In this respect, it should also be noted that what might be considered the most intimate relationship of all, that between spouses, does not provide us with a distribution of forms of address strikingly different from that in other types of relationship. The majority of forms of address appear to be epistolary (60%). In this case, we have to take into account possible differing power relations. It is not self-evident that Early
Modern husbands and wives should be treated as equals, and in case studies on English, for instance, the category of spouses has been broken down into that of husbands and that of wives, with significant results (e.g. Nevala 2004b). In the Northern Netherlands, unsurprisingly, Early Modern society was patriarchally organised, and although this certainly does not imply that the world of women was always bound by the walls of their homes, it seems reasonable to assume that, in general, women were subordinate to men. Still, there are indications that the balance of power may have been more equal within families than in the public sphere (Frijhoff & Spies 1999: 190–192). Furthermore, Nevala (2004b: 173–178) argued that in the case of English correspondence, an increase in positive politeness could be witnessed from the fifteenth to the seventeenth century, indicating differences in the power status of wives who became more equal to their husbands than previously, and thus came to use thou more often. To clarify any possible links between power relations and politeness strategies in our letters, we also divided the data for spouses into those for husbands and those for wives. The results are given in Table 4.

From Table 3 above, it could be gathered that spouses used epistolary forms in 60% of all cases. In Table 4, we see that husbands writing to their wives score below this average, using 55% epistolary forms, whereas wives writing to their husbands score somewhat higher with 67% epistolary forms. Thus a slight change appears in the direction of more negative politeness in letters by supposed inferiors to supposed superiors, mirroring the lower proportion of epistolary forms in letters from parents to children, that is, from superiors to inferiors. There are two caveats, however. First, and as before, it should be noted that even if there is a power-related difference in the use of negatively polite forms, the converse to this is that husbands writing to their wives still use epistolary forms in more than half of all cases. Put differently, the principal politeness strategy employed by these supposed superiors writing to their supposed inferiors would be a negative one. Secondly, the informal or spoken-language j-forms appear in

<table>
<thead>
<tr>
<th>Spouses 17th century (autographs)</th>
<th>Husband &gt; wife 51 letters 44 writers</th>
<th>Wife &gt; husband 48 letters 38 writers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ul (nom/obl/poss)</td>
<td>465</td>
<td>53</td>
</tr>
<tr>
<td>ue (nom/obl/poss)</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Subtotal epistolary forms</td>
<td>487</td>
<td>55</td>
</tr>
<tr>
<td>gij (nom)</td>
<td>193</td>
<td>22</td>
</tr>
<tr>
<td>u (obl/poss)</td>
<td>201</td>
<td>23</td>
</tr>
<tr>
<td>j-form (nom/obl/poss)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>u (nom)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Total</td>
<td>881</td>
<td>100</td>
</tr>
</tbody>
</table>
letters by wives. In Table 3, we saw that the highest proportion of j-forms occurs in letters by parents writing to children, and even though numbers are low for j-forms in general, this would seem in line with the balance of power. Here, however, all j-forms in letters between spouses turn out to be produced not by the superiors, that is, by the husbands, but by the inferiors, namely the wives, and by three female letter writers in particular. How can this be explained?

As we have argued in Chapters 4 and 5, and will argue in Chapters 7 and 8, writing experience is an important variable in historical sociolinguistics. As a rule, j-forms were not used in written language, and in those cases where they were used, this was mainly in representations of (lower-class) spoken language in farces and such like. We have also argued that there were important differences in writing experience between men and women, as well as between the different social ranks. The j-tokens could, therefore, be explained as a result of women's lesser writing experience, that is, of less familiarity with the conventions of the written code. But this would obviously be at odds with women's principal (negative) politeness strategy, with no less than 67% epistolary forms when writing to their husbands. How can informal, spoken language features such as j-forms, which must be positively polite if evaluated in terms of politeness strategies, appear in letters where the main politeness strategy is a negative one? Note, furthermore, that female j-users do not confine themselves to j-forms. Instead, we find mixed usage such as in (1), where the terms of address are in boldface.

(1) want ick seer verlangh om van ul eens schrivens te hebbe
because I very long to from you one letter to have
want ick suent den drieden mei geen schrivens van ul gesien
because I since the third May no letters from you seen
en hebt en ick oock seer verlan na tingh van ul te hebbe
NEG have and I also very long for news from you to have
o doet toch sooveel en schrift eens meten erste soo dree
oh do+IMP indeed so-much and write+IMP once with-the first as soon
als ghi kent hoe het daer al met ul is en hoe het daer al
as you can how it there all with you is and how it there all
gaet en of ie den toback al verkoopt
goes and if you the tobacco already sell
‘because I very much long to having a letter from you, because I have not seen any letters from you since 3 May, and I also very much long for news from you, oh with the first [occasion] as soon as you can, please write how you are doing and how things are going over there and if you are already selling the tobacco’

In this excerpt from a 1664 letter to her husband Gerrit Ariaanse van den Broek, Maartje Jaspers most often uses the epistolary form ul, as she does in the entire letter. In (1), she switches to ghi when she needs a nominative form, then uses ul again following the preposition met “with”, and then produces her single j-form, viz. ie, that is,
je, the reduced form of jij “you”. The passage is overtly affective: the writer is longing for news about her husband’s well-being. There is no transition to even more affectivity warranting the switch to the informal, positively polite j-form. In fact, it may even be argued that the final clause in (1), which refers to her husband’s commercial success, is less affective than the preceding clauses. This mixing of so-called positively and negatively polite strategies calls for an explanation, and as we will argue in Section 3.5, this does not necessarily – or in any case not primarily – involve an appeal to politeness. In another effort to come to terms with politeness strategies, we will focus specifically on the j-forms, as these are the most obvious informal and positively polite variants.

3.3 The j-forms

In the previous sections, we mainly focused on the proportion of prototypically epistolary forms relative to the other terms of address, on the assumption that epistolary forms such as ul and ue index negative politeness. Another way of investigating possible politeness strategies is by focusing on the unambiguous T-forms. As stated above (Section 2 and Section 3.2), the j-forms probably represented the most informal register, were largely restricted to spoken language, and thus may have indexed positive politeness. In this section, we therefore concentrate on the use of j-forms in both the seventeenth- (3.3.1) and the eighteenth-century subcorpus (3.3.2).

It should be noted in advance that the j-pronouns, while being T-forms in present-day Dutch, are not unanimously considered to have indexed informality in earlier stages of the language such as in the seventeenth century (Vermaas 2002: 47; Aalberse 2009: 56). In fact, they have been found to vary with gij and u in seventeenth-century texts, in both polite and intimate contexts (Vermaas 2002: 48), thus suggesting interchangeability. This is one of the reasons why Aalberse (2009: 55) refuses to specify the politeness level of j-forms as either T or V. She also indicates, however, that j-forms were associated with the spoken language and were used in informal registers (2009: 58), which would qualify them as T-forms, according to her own description of T-forms as “informal address forms”, as opposed to V-forms as “polite address forms” (Aalberse 2009: 52). Considering the fact that j-forms have been infrequent in writing well into the nineteenth century, and were moreover mainly used in informal settings in plays in the seventeenth and eighteenth centuries, we deem it safe to regard them as more intimate and more informal than gij, u and epistolary forms of address, and thus relative to the latter forms, as T-forms.

3.3.1 The seventeenth century

The seventeenth-century subcorpus of autograph private letters, which contains 2,776 terms of address, has only 56 j-forms, amounting to 2%. This is in line with the results presented in the previous sections. These 56 tokens are produced by 11 different letter writers. Of the 56 tokens, 29 are nominatives, mostly the reduced variant je, and 27
j-forms are obliques or possessives. The predominance of the reduced nominative *je* (28 tokens) as opposed to only one token of the full form *ijj* demonstrates the oral origin of the j-forms. As in present-day Dutch, it must have been the reduced forms that were common in unstressed positions. In the absence of a writing tradition incorporating j-forms, the letter writers put the reduced spoken language form onto paper, whereas currently the full forms are conventionalised in written Dutch, even when reduced forms would be used in the spoken language (ANS 5.2.7).

In Section 3.2, we discussed an excerpt from a 1664 letter containing both *ul*, *gij* and *je* (in normalised spelling). In order to gain a fuller understanding of the variation, we will now discuss three more letter writers who used j-forms (out of 11 users of j-forms in total), and see whether we can explain their occurrence in terms of politeness strategies. In 1664, Leendert Janss wrote a letter to his son Jan Leenders. In this letter, we find a mixed usage of forms of address. The four tokens of *ul*, two in the first part, and two in the closing part, all appear in epistolary formulae (see Chapters 3–5) such as the greeting formula in (2) and the subscription in (3).

(2) *wenssen Ul hondert duissent goed nacht*
    wish     you hundred thousand good night
    ‘[I] wish you a hundred thousand good nights’

(3) *VI vader Leendert Janss*
    your father Leendert Janss

In between, in the less formulaic middle part of the letter, Leendert Janss mixes the nominative *gij* with the possessive form *jou*, as in (4).

(4) *past dat ghy nyet verghet teschryuen mijt*
    watch that you not forget to-write with
    *die eersten scheppen want jou moeder hart verlanckt*
    the first ships because your mother very longs
    *hoe dat ghy alvaert*
    how that you all-fare
    ‘mind you do not forget to write with the first ships for your mother very much longs [to hear] how you are doing’

A similar distribution is found in a letter by Cornelis Cornelisz van de Stadt, written to his wife Maerij Pietersz in 1663. Again, *ul* is mainly found in formulae (5), with the nominative *gij* and a possessive *j*-form appearing in the middle part, but then also an oblique *j*-form in a formulaic sequence in the letter’s ending (6).

(5) *jck laet ul weten als dat jck noch gesont ben jck hoop dat het met ul*
    I let you know as that I still healthy am I hope that it with you
    *en Vader en Moeder oock soo is*
    and father and mother also so is
    ‘I let you know that I am still healthy. I hope it is the same with you and father and mother’
Chapter 6. Forms of address

(6) jck bidd jouw mijn lieve
    I beg you my dear
Vrouw schrijft doch
    wife write+IMP indeed
‘I beg you, my dear wife, please write’

It may be significant that the j-forms are restricted to oblique and possessive use with these authors. If these were preferred to u, the historical oblique and possessive form in the gij-paradigm (Section 2), then this could very well be a positive politeness strategy, possibly to avoid u for its formal resemblance to the epistolary forms. At the same time, however, gij is used instead of nominative j-forms. If the j-forms were informal or colloquial or the most positively polite alternatives, which seems reasonable to assume, then using gij would create relative deference.

There are letters where the mixing of forms is even less clear-cut, and where nominative j-forms do in fact appear. One telling example is the letter that Trjntie Batens wrote to her husband Hendrick Batens in 1664, which contains no less than 25 j-forms. She also uses the epistolary form ul in formulaic language at the beginning and at the end of the letter. Moreover, when using a discourse-continuing formula in the non-formulaic middle part, she also draws on ul (7–9).

(7) voorts laat ick ul weten als dat die persoon
    further let I you know as that that person
    die u de 3 rijckxdaelders geleent heeft bij mij geweest is
    who you the three rix-dollars lent has to me come is
‘further I let you know that that person who lent you the three rix-dollars came to me’

(8) ick laet ul noch weten als dat
    I let you yet know as that
    ons jongste kindt henderick doot is
    our youngest child Henderick dead is
‘I also let you know that our youngest child Henderick is dead’

(9) voorts mijn alderliest laet
    further my dearest let
    ick ul weten als dat hier de spraeck gaet van de engelse oorlogh
    I you know as that here the talk goes about the English war
‘further, my dearest, I let you know that there are rumours about the English war’

Throughout her letter, however, in the non-formulaic middle part, we find nominative j-forms, nominative gij, and oblique and possessive uses of u. In (10), the writer switches from the reduced nominative form je to the possessive and oblique form u.
maer ije moet mij
wel veracht hebben maer nu wert hier veel gepraet van u
well despised have but now is here much talked about your
swager dat die u menigmael verleijt heeft
brother-in-law that he you many times seduced has
‘but you must have despised me, but now people talk about your brother-in-
law, that he has seduced you many times’

The mixing of forms in this letter is surprising when compared with the previous two letters. In those letters, nominative gij was mixed with oblique and possessive j-forms. Here, nominative j-forms are mixed with oblique and possessive u. If the writer chose j-nominatives so as to strengthen the intimacy existing between her and her husband, then the many instances of the presumably more distant form u stand out. The whole letter is written in a fiercely affectionate style, related to the severe troubles the writer finds herself in, such as financial problems and social exclusion. This could, of course, explain the many colloquial j-forms, but as with the previous two letters, this leaves unexplained why we also find u, which would be relatively distant or create deference from the perspective of the j-forms.

Similar to letter writer Maartje Jaspers in Section 3.2, the three writers discussed here vary in their use of forms of address, mixing j-forms with forms from the gij/u-paradigm as well as epistolary forms. Single j-form users do not exist. There may have been a division of labour between epistolary forms and pronominal forms, at least for some writers, who largely restricted the use of epistolary forms to formulaic stretches of discourse, and the use of pronominal forms to non-formulaic parts. When focusing on these non-formulaic parts, there are still no single j-form users to be found. We may wonder what the influence of politeness strategies could be. Pronominal forms may have been considered awkward in epistolary formulae (but see (6)), which could be a negative politeness strategy, but which could also result from the fact that formulae were probably learnt by heart and memorised as a whole, including the forms of address commonly used in the written language (Chapter 4). Furthermore, if the use of pronominal forms in the non-formulaic middle part of letters and – in particular – the absence or avoidance of epistolary forms indexed positive politeness strategies, including avoiding negative politeness, then an explanation is needed for why negative politeness strategies were so common in the beginnings and endings of letters. Besides, many letters writers did in fact use ul in non-formulaic sequences, as will be explained below (Section 4.2.4). Finally, the variation of informal or colloquial or positively polite j-forms and gij/u-forms needs to be explained, and the letters at hand do not provide us with many clues in terms of negative or positive politeness strategies. We established the variation, but its pragmatic significance remained obscure.

Meanwhile, we should recall the very low proportion of the unambiguously positively polite j-forms, in particular when considering the fact that most of our letters
were written to close family members. In the middle part of their letters, the writers discussed here acted as many other letter writers did, namely using \textit{gij} and \textit{u} (see also Section 4.2.4 below). In addition, they incorporated \textit{j}-forms, which must have been common in the spoken language, but which are marginal in our corpora. In our view, what we are witnessing here is very similar to what we observed in Chapter 2 on the interplay of local phonology and supralocal writing systems. Although there were supralocal conventions, writers occasionally diverged from these conventions, admitting straightforwardly oral features into their writing. In other words, \textit{gij} and \textit{u} bear witness to the fact that the writers aimed to use a supraregional variety, with occasional interference of local features such as \textit{j}-forms.

In such an interpretation, \textit{j}-forms would not be considered as markers of positive politeness, but as traces of the spoken language. As a result, we would expect \textit{j}-forms to appear mainly in letters written by less experienced writers, which means more in letters from the lower and lower middle classes than in letters from the upper and upper middle classes, and more in letters by women than in letters by men (cf. Chapter 4). Even though numbers are low, since there are only 56 tokens of \textit{j}-forms, such a distribution according to writing experience appears to be borne out by our results. In Table 5, we give the proportion of \textit{j}-forms across social class in the seventeenth-century autograph corpus.

Whereas the \textit{j}-forms take up about 5\% of all forms of address in the lower (middle) classes, their proportion is approximately 1\% in the UMC, with not a single token in the UC. The LC and LMC tokens are produced by eight writers, and the UMC tokens by four writers. Comparing these numbers to the total number of writers across social class in the seventeenth-century autograph subcorpus again confirms that the LC and LMC are overrepresented. Among the 46 LC and LMC writers in the seventeenth-century autograph subcorpus, eight of them use at least one \textit{j}-form (18\%), whereas four \textit{j}-form users (4\%) are found among the 105 UMC writers. Inclusion of the UC (17 writers) would reduce the proportion of \textit{j}-form users in the upper (middle) classes even more. As to gender, the six men and five women among the 11 writers using at least one \textit{j}-form indicate an overrepresentation of women since the total number

<table>
<thead>
<tr>
<th>17th century (autographs)</th>
<th>Terms of address</th>
<th>\textit{j}-forms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>\textit{j}-forms</td>
</tr>
<tr>
<td>LC</td>
<td>150</td>
<td>8</td>
</tr>
<tr>
<td>LMC</td>
<td>706</td>
<td>35</td>
</tr>
<tr>
<td>UMC</td>
<td>1,628</td>
<td>13</td>
</tr>
<tr>
<td>UC</td>
<td>292</td>
<td>–</td>
</tr>
<tr>
<td>Total</td>
<td>2,776</td>
<td>56</td>
</tr>
</tbody>
</table>
of male and female writers in the autograph subcorpus (219 letters by 168 writers) is 122 and 46, respectively.

As stated previously, numbers are low. Nevertheless, it seems that the lower and lower-middle classes and women are overrepresented in the results for $j$-forms. This is in line with the expectations, if indeed $j$-forms are traces of the spoken language occurring mainly in the written language of less experienced writers. We will now turn to the eighteenth-century subcorpus to see whether the distribution of $j$-forms there might provide us with more clues for their interpretation.

3.3.2 The eighteenth century
The picture of the eighteenth-century results does not differ greatly from that which became apparent from the seventeenth-century results, except in one respect, to which we will turn below. First, it should be noted that the token count of $j$-forms attested in the eighteenth-century subcorpus is even lower than the 56 tokens found in the seventeenth-century subcorpus. Whereas the total number of terms of address is about twice as high in the eighteenth-century subcorpus, viz. 5692, the number of $j$-forms drops to no more than 37 (1%, see also Sections 3.4 and 4.1 below). If $j$-forms were markers of positive politeness, this would imply a reduction of positive politeness from the seventeenth to the eighteenth century. If $j$-forms represent traces of the spoken language, this would imply a reduction of oral interferences. We will argue that $j$-forms should be considered as oral features which occasionally appear in the written language of less experienced writers.

As in the previous century, we do not find letter writers in the eighteenth-century subcorpus who only use $j$-forms. All $j$-form users show mixed usage of forms of address, mostly combining $j$-forms with $gij$ and $ue$. By way of illustration, we will focus on one of them, viz. Steven Cadenski, an upper-class adolescent from Sri Lanka who leaves for the Netherlands in the 1780s to study theology. In 1781, he was at the Cape of Good Hope in southern Africa, from where he wrote several letters to family members in Colombo, Sri Lanka, such as his siblings, whom he advised on how to behave. The first excerpt (11) is from a letter to his sister Mietje, dated 30 May 1781.

(11) dat uwEd: zeer
that you very
voorzichtig moet zijn, in al uw handel
careful must be in all your behaviour
en wandel, en wagt altoos voor een kwaade
and behaviour and mind always for a bad
naam, vermijd kwaad gezelschap: [...]name avoid bad company
[...] neem dan in alles, onze
take than in all our
deugdzame zussie Stijntje tot voorbeeld,
virtuous sister Stijntje as model
hoewel uwEd: wat kinderagtig is, hoope although you somewhat childlike are hope
ik egter wèl, dat uwEd: zoo wijs zal zijn, I however well that you so wise will be
dat uwEd: zich niet in eenig gevaar van kwade that you refl not in any danger of bad
name zal steeken
name will put
‘that you should be very careful, in all your behaviour, and always mind you do not get a bad name. Avoid bad company [...] In every respect, use our virtuous sister Stijntje as a model. Although you are somewhat childlike, I do hope that you will be so wise you will not put yourself in the danger of a bad reputation’

The excerpt in (11) contains four nominatives uwEd: and a possessive form uw, which historically belongs to the paradigm of gij (Section 2). Cadenski’s didactic tone, and his struggle with the family honour, do not lead him to the use of j-forms. In terms of politeness, one could argue that such explicit advice would be rendered more convincing when verbalised in a fairly familiar, positively polite manner, which would decrease distance. There are, however, no signs of such a politeness strategy. One could also argue that such harsh words need to be rendered in negatively polite words, so as to cancel the face-threatening act. About two weeks later, however, on 18 June 1781, Cadenksni wrote a similar letter with advice to his brother Sibo in Colombo, Sri Lanka; see the excerpt in (12).

(12) leef nu niet meer als voor
live now not anymore as before
deezen, gelyk Leeuwen en Beeren, maar
this like lions and bears but
als menschen, als reedelijke Schepzels
as humans as rational creatures
die vernuft hebben, onder doet altemet
who reason have under do+IMP always
uw broeder Bram, wijl hij dog jou
your brother Bram because he indeed your
ouder broeder is, daar bij zal je niets
older brother is there by will you nothing
verliezen, en verzoent met Hem, zoo het lose and appease with him so it
tot als nog niet geschied is, ik verzoeken
until as yet not happened is I beg

4. We consider uwEd: as a variant of ue. See Section 4.2.2 for discussion.
uwEd: zoo ras als uwEd: dit leest en  
you as soon as you this read and  
mijn sentiment verstaan zult hebben,  
my opinion understood will have  
moet gij zonder uitstel met Hem verzoenen  
must you without delay with him appease  
‘do not live as before, as lions and bears, but as humans, as rational creatures  
who have the power of reason. Always obey your brother Bram, because he is  
your elder brother after all. You will lose nothing by it. And appease him, if it  
has not happened until now. I beg you, as soon as you read this, and will have  
understood my opinion, you must without further delay appease him’

In (12), Cadenski uses uwEd: both as a nominative and as an oblique, as well as a  
nominative and a possessive j-form, and also nominative gij and possessive uwe.  
Clearly, Cadenski, who prefers ue or variants of ue in these and other letters contained  
in our corpus, may have switched to more positively polite alternatives such as gij and  
uwe, and even two j-forms, possibly because he wanted to decrease distance in this  
overly consultative letter. But then it remains to be explained why he adhered to ue in  
the letter to his sister. Could we acquire a better understanding of this variation by ap-  
proaching it from a different perspective, i.e. the oral-written perspective that we took  
in the previous section?

If the sporadic occurrence of j-forms can be explained as oral interferences in the  
written language, we would expect most of the tokens to appear in letters by less expe-  
rienced writers. Therefore, as before, we have looked at the distribution of j-forms across  
the social classes in the eighteenth-century subcorpus. Table 6 presents the results. The  
eighteenth-century use of j-forms does not show any social differences, but what is  
striking, is the overall decrease of j-forms. Whereas the LC and LMC scored 5% or  
more in the seventeenth century, they seem to converge to the use of the UMC and UC  
in the eighteenth century, which implies using hardly any j-forms. Interestingly, the

<table>
<thead>
<tr>
<th>18th century Terms of address</th>
<th>Total</th>
<th>j-forms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>LC</td>
<td>327</td>
<td>4</td>
</tr>
<tr>
<td>LMC</td>
<td>1,096</td>
<td>6</td>
</tr>
<tr>
<td>UMC</td>
<td>1,746</td>
<td>6</td>
</tr>
<tr>
<td>UC</td>
<td>2,523</td>
<td>21</td>
</tr>
<tr>
<td>Total</td>
<td>5,692</td>
<td>37</td>
</tr>
</tbody>
</table>
UC does now in fact provide us with \( j \)-forms, as opposed to the seventeenth-century results. Taking into account which writers were responsible for the 37 attested \( j \)-forms, we observe that 16 writers produced 37 tokens. Of these 16 writers, 5 are allocated to the LC and LMC and 11 to the UC and UMC. Even if we compare these numbers to the total number of writers, with 191 writers in the UC and UMC and 101 writers in the LC and LMC, this still means that the UC and UMC are equally well represented as the LC and LMC. This may come as a surprise, since it counters the assumption that the users of \( j \)-forms were mainly less experienced writers. On closer inspection, however, it turns out that five of the 16 writers using \( j \)-forms received the regional code “abroad”, which means that they were born and/or raised abroad in Dutch-speaking families. These five writers produced 17 \( j \)-forms, thus almost half of all eighteenth-century \( j \)-forms. Four of these writers are UC men, and one of them is Steven Cadenksi (see (11) and (12)). If we were to leave out people born and raised outside the Netherlands, the same generalisations as in Section 3.3.1 would apply: women produced more \( j \)-forms than men, and the LC and LMC produced more \( j \)-forms than the UC and UMC.

As argued in Chapter 4 and 5, writing experience was at least in part determined by social and gender factors. We conceived of writing experience as the degree to which a person is connected to the written culture. From the results in this section, we conclude, that being born and/or raised abroad relatively disconnected people from the written culture as it developed in the Netherlands, and thus from conventionalised usage. Naturally, this will also apply to the spoken language. The lack of letters from “abroad” means we are not in a position to fully investigate possible linguistic differences between “home” and “abroad”, but the matter brings to mind similar cases such as the maintenance of conservative forms in eighteenth- and nineteenth-century Canadian French as reported by Martineau (2007). It also recalls language attrition as found in the language of, for instance, twentieth-century Dutch migrants to Canada (Keijzer 2007). Such cases are not exactly similar, but what unites them is the fact that the language of migrants displays forms diverging from conventional usage in the home country due to the speakers’ relative isolation from the original language community.

Finally, we have identified a number of tendencies, such as that women are more likely to use \( j \)-forms than men, and likewise LC and LMC writers more than UC and UMC writers, and people from abroad more than people born and raised in the Netherlands. These tendencies clearly point to \( j \)-forms as oral interferences in the language of less experienced writers. At the same time, we should recall that the overall number

5. Note that we are discussing letters and letter writers with the regional code “abroad”. As indicated in Chapter 1 (Section 5), this refers to writers born and/or raised in Dutch speaking families outside the Netherlands (“second generation”). This category does not include letters from writers born and raised in the Netherlands, who had migrated abroad, either temporarily or definitively (“first generation”).

of j-forms is extremely low in the eighteenth century, as it was in the seventeenth century. This is a reminder that writing experience is a gradual notion, and that less writing experience may also mean that a writer only sporadically diverges from very strong general patterns of conventional usage. In the following sections, we will focus on these strong general patterns, so as to evaluate the explanatory power of politeness strategies once again.

3.4 An increase in negative politeness?

As stated in Section 2 above, there are very clear patterns in the use of terms of address in both the seventeenth and the eighteenth century. We will detail these patterns in Section 4, where we report on the sociolinguistic distribution of forms in our subcorpora. Here, we anticipate some of the results in order to relate them to the discussion of politeness developed in the previous sections. Despite the wide range of pronominal variation that has existed in the spoken language of the seventeenth and eighteenth centuries, the written language is largely restricted to a small number of forms. In the seventeenth century, the four dominant variants are gij, ul, ue and u (oblique and possessive). In the eighteenth century, three variants, viz. gij, ue and u (oblique and possessive) prevail. In Table 7, we present the overall results of the main variants in both absolute numbers and percentages.

Focusing for the moment on the epistolary forms, two observations can immediately be made. First, the most frequent seventeenth-century variant, viz. ul (42%), is reduced to marginality in the eighteenth century (1%). Secondly, the total score for the epistolary forms ul and ue nevertheless rises from 58% to 72%, which is due to the sharp increase of ue. In terms of politeness strategies, this would mean a diachronic increase in negative politeness.

Such an increase in negative politeness seems surprising when considering, for instance, the contrasting increase in positive politeness strategies established for Early Modern English letter writing (Nevalainen & Raumolin-Brunberg 1995; Nevala 2003, 2004a) and the observed development towards simpler forms of address (i.e. reduction of negatively polite forms) in late-eighteenth- and early nineteenth-century Dutch

<table>
<thead>
<tr>
<th></th>
<th>Forms of address, overall results</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ul</td>
</tr>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>17th century</td>
<td>1,172</td>
</tr>
<tr>
<td>Subtotal epistolary forms: N=1,606 (58%)</td>
<td></td>
</tr>
<tr>
<td>18th century</td>
<td>53</td>
</tr>
<tr>
<td>Subtotal epistolary form: N=4,110 (72%)</td>
<td></td>
</tr>
</tbody>
</table>
elite correspondence (Ruberg 2005). We could possibly come to terms with these opposite developments, if we attribute them to different times and to different social levels. The increase in positive politeness strategies in English mainly characterises the sixteenth and seventeenth centuries; the “politeness context” could be different for the Dutch letter writing of the seventeenth and eighteenth centuries. A time and social level difference could be a factor in the Dutch elite letters that, in the course of the nineteenth century, showed a decrease of lengthy and “humble” forms of address and a dwindling use of *ue* in favour of *gij, u, j*-forms and the French alternative *tu* (Ruberg 2005: 73). It should be noted that our corpora do not comprise letters from the highest level of the nobility and non-noble ruling class (see Chapter 1, Section 5).

If epistolary terms of address signal negative politeness, then their increase from the seventeenth to the eighteenth century remains remarkable, not simply because of the seemingly contrasting developments discussed. The diachronic increase in epistolary forms as such remains an unexplained phenomenon and their usage alongside other forms of address which implies a supposed mixture of positive and negative politeness strategies also raises more questions than the politeness approach answers. In the next section, we will, therefore, suggest a different approach.

### 3.5 Politic behaviour and social stratification

In the previous sections, we repeatedly attempted to interpret the use of forms of address in our subcorpora from the perspective of politeness strategies. Comparing private letters with business letters (Section 3.1), we observed a remarkable difference in the use of epistolary forms, which suggested that these were, at least relatively, negatively polite. Similar conclusions could be drawn when comparing the forms of address used within the core family (3.2). We also noticed, however, that epistolary forms stood out anywhere, in any context, on average taking up 50 to 60% of all forms of address, and their share even increased to over 70% in the eighteenth century (3.4). At the same time, unambiguous T-forms such as *j*-pronouns were marginal in both the seventeenth and the eighteenth century (3.3). And while it is possible to conceive of *gij* and *u* as relatively positively polite, that is, when compared to epistolary forms (cf. Bax 2010: 53–55), then the supposedly positive forms are still outnumbered by the supposedly negative forms in both centuries (3.4). It should be noted, in this respect, that *gij* and *u* have been the dominant variants in published texts, including grammar books.

---

6. One example referred to by Nevala (2003) is the rise of possessive pronouns such as *my* in forms of address, which would mark positive politeness. Recall that we established the opposite development in this respect as well, with the possessive pronoun almost entirely vanishing from usage in the eighteenth century (Chapter 3, Section 6.1.3).

7. Note, in this respect, Kádár & Culpeper’s (2010: 18–19) principled comment that, for obvious reasons, we do not tend to have hearer’s evaluations of polite or impolite behaviour in historical documents.
well into the nineteenth century (cf. Section 2). Summing up, labelling forms of address as inherently positively or negatively polite did gain us some insight, but particularly the sheer number of epistolary forms in itself calls for another approach.

It has been remarked previously that Brown and Levinson’s approach can only distinguish between polite and impolite behaviour which does not leave any room for an unmarked middle ground (Jucker & Taavitsainen 2003: 11). For such unmarked speech situations, where the speaker is neither polite nor impolite, and does not use overtly negatively polite forms nor overtly positively polite forms, Watts (1989, 2003) has proposed the term politic behaviour: “behaviour, linguistic and non-linguistic, which the participants construct as being appropriate to the ongoing social interaction” (2003: 20). If the social situation changes, when the topic is professional rather than personal, for instance, or when the hearer is superordinate rather than equal or subordinate, then the speaker may change his behaviour. In such cases, the speaker may use terms which are seemingly more polite than the terms he used before, which does not mean, however, that they are more polite in the new situation: “there are linguistic structures in excess of what the speaker needs to utter which nevertheless go unnoticed, since they form part of the reproduction of institutionalised discursive formats” (Watts 2003: 19). As the situation changes, the language may change, without becoming more polite or impolite, but simply remaining politic in a different context. Similar comments have been made by Meier (1995: 351), who noted “that it is not different types of politeness that are called for, but rather a return to focussing on how certain linguistic features (not dubbed inherently polite or impolite) pattern and are perceived in particular contexts to fulfil certain functions. This is a quest to explore socially appropriate language behavior”. Only exceeding this politic or socially appropriate language behaviour, either positively or negatively, leads to behaviour that is polite or impolite behaviour, or over polite or over impolite.8

Bax (2010), focusing on seventeenth-century Dutch elite correspondence, has established changes in politic behaviour in accordance with changes in the communicative situation, without assuming any type of behaviour being more polite or impolite than another. We will take the same approach here, and assume that the forms of address found in our letter corpora signal neither politeness nor impoliteness, but should in general be considered the appropriate linguistic behaviour in private correspondence.

8. Watts (2003: 21) stresses that “Polite behaviour will therefore be behaviour beyond what is perceived to be appropriate to the ongoing social interaction, which says nothing about how members evaluate it”. Cf. also Meier (1995: 352): “Certain address forms (e.g. sir), for example, would not be described as more polite than others, but would convey a certain amount of deference within a particular speech community. Over politeness could be a case of an excessive amount of deference being shown in a situation that warrants less. A form associated with a high degree of deference could thus be inappropriate as well as appropriate, depending on interlocutors’ perceptions of a particular situation. This is also true for routine formulae (e.g. greetings) and lexical items (e.g. please, thank you). All are expected in certain contexts but also can occur inappropriately in other contexts.”
Given that the majority of forms of address are epistolary forms such as *ul* and *ue*, we refrain from considering these as overtly or negatively polite, and instead consider them as conventional usage, that is, as politic behaviour.

This opens up a different perspective. If we refrain from judgments on politeness or impoliteness, the variants presented and discussed in the previous sections may indeed be variants in the sociolinguistic meaning of the word. Perhaps, they do not carry in themselves different meanings such as “more polite or impolite”, but could be considered as neutral variants of the sociolinguistic variable “form of address”. This would allow a variationist sociolinguistic investigation using variables such as gender and social class. In this respect, it should be noted that van Leuvensteijn (2000a) studied the use of terms of address in the eighteenth-century epistolary novel *Historie van Mejufrouw Sara Burgerhart*, co-authored by Betje Wolff and Aagje Deken and first published in 1782. He concluded that the use of forms of address could not be explained with reference to politeness models, but as patterned with the social class index. The lower classes preferred *j*-forms and the upper classes preferred *gij* and *u*. Even though this usage does not conform to what we find in our letters, it is revealing that external sociolinguistic variables proved more reliable than an appeal to politeness strategies.

Returning, for a moment, to the data presented in the previous sections, there are several clues for such an approach. Until now, we have treated *ul* and *ue* as one category of epistolary forms. It should be recalled, however, from Section 2 that the WNT claims that *ul* was in use in the fifteenth, sixteenth and seventeenth centuries, and *ue* from the sixteenth to the nineteenth century. We have also seen in Section 3.4 that *ue* was much less frequent than *ul* in the seventeenth century, while *ul* was almost absent from the eighteenth century and *ue* had become the dominant variant. What appears to be going on here is the well-known phenomenon of an incoming variant replacing an older variant, and this obviously calls for sociolinguistic analysis. Finally, when discussing the sporadic *j*-forms in Section 3.3, we have already alluded to a possibly sociolinguistic interpretation, notably in terms of writing experience, which we consider of central importance to our understanding of the written language in historical private letters. In the next section, we will report on case studies from such a sociolinguistic perspective.

4. Sociolinguistic variation and change

In this section, we will discuss the use of forms of address in our seventeenth- and eighteenth-century subcorpora from a variationist point of view. We will first present the overall results (4.1) before turning to the epistolary forms (4.2) and the pronominal forms (4.3).

---

9. See Section 4, and see Simons (2013:134–138) for a detailed discussion of the differences between the results from our letter corpora and the evidence from literary sources presented by van Leuvensteijn (2000a, 2000b).
4.1 The overall diachronic picture

Figure 1, which visualises the results of Table 7 in Section 3.4, presents the overall picture of the main variants of forms of address, for both the seventeenth and the eighteenth century. As we have seen before, the main seventeenth-century variants are *ul*, *ue*, *gij* and *u* and the eighteenth-century ones are *ue*, *gij* and *u*. Note that *u*, which is originally the oblique and possessive form of *gij*, does not co-occur only with nominative *gij*. Nobels (2013: 77–114) shows that seventeenth-century letter writers do not choose between either epistolary forms or the historical *gij/u* system, but may vary widely. Letter writers who use only *gij* in nominatives, may use either *u* in the oblique and possessives, or epistolary forms, or a combination of *u* and epistolary forms, the last option representing the dominant strategy by far. This implies that we cannot combine *gij* and *u*, but rather that we need to distinguish them, even though they are functionally different from the epistolary forms that are used in both nominative and oblique or possessives.

As argued above, *ue* is the incoming variant, which first competes with *ul* (in the seventeenth century) and then prevails (in the eighteenth century), relegating *gij* and, to a lesser extent, also *u*. The focusing of a formerly more varied linguistic situation is exemplified by the high grey column representing the proportion of *ue* in the eighteenth century. Variation is levelled out to a great extent, and *ue* emerges as the preferred variant.
4.2 Epistolary forms of address

In this section, we discuss the distribution of the epistolary forms of address, both in the seventeenth (4.2.1) and in the eighteenth century (4.2.2), before summarising the diachronic development in 4.2.3. One diachronic result does not seem to be in line with the general development, to which we will turn in Section 4.2.4.

4.2.1 The seventeenth century

In the seventeenth century, 1,606 out of 2,776 forms of address were epistolary, which amounts to 58% (cf. Table 7 in Section 3.4). In Table 8, the epistolary forms of address are divided according to social class.

While the overall proportion of epistolary forms is 58%, the proportion by social class varies from slightly less to slightly more: the LMC, for instance, scores 48%, while the UMC peaks at 64%. We do not wish to attribute any great significance to these differences. In Section 3, the proportion of epistolary forms of address was approximately 50 to 60% in almost any selection of letters made, and we believe the results in Table 8 corroborate this. Generally speaking, epistolary forms take up 50 to 60% in all social classes, and we will show below that they tend to appear somewhat more often in the upper classes than in the lower classes. This is what we see in Table 8 when comparing the UMC to the LC and LMC. Note, however, that the UC seems to slightly lag behind when compared to the UMC. We will return to this in Section 4.3.

As argued in Sections 3.5 and 4.1, we should distinguish between the two epistolary forms, \( ul \) representing older usage, and \( ue \) being the incoming variant. Splitting up the results into proportions of \( ul \) and \( ue \) across social class generates interesting results, as can be seen in Figure 2. In the LC, all attested epistolary forms are tokens of \( ul \). In the LMC, the proportion of \( ul \) drops only slightly to 96%. For both social ranks, therefore, we conclude that if an epistolary form is used, it is highly probable that it is \( ul \). There is a clear difference with the UMC and the UC, where \( ul \) makes up only about 60% of all epistolary forms, and where, consequently, \( ue \) rises to 35% and 40%, respectively. The rising variant, in other words, comes in via the UC and UMC.

Table 8. Epistolary forms of address, by social class.

<table>
<thead>
<tr>
<th>Epistolary terms of address</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>( N )</td>
<td>%</td>
</tr>
<tr>
<td>LC</td>
<td>80</td>
</tr>
<tr>
<td>LMC</td>
<td>338</td>
</tr>
<tr>
<td>UMC</td>
<td>1,042</td>
</tr>
<tr>
<td>UC</td>
<td>146</td>
</tr>
<tr>
<td>Total</td>
<td>1,606</td>
</tr>
</tbody>
</table>
As there is some intra-speaker variation in the use of epistolary forms (Section 3.2, 3.3, also 4.2.4 below), it is interesting to investigate the number of letter writers who use *ul* and *ue* at least once in their letter(s). The results of such an analysis confirm that *ue* entered through the UC and UMC. As can be seen in Figure 3 below, every letter writer (100%) assigned to LC uses the epistolary form *ul* at least once. The proportion of writers using *ul* at least once drops steadily to 83% in the LMC, to 75% in the UMC, and to only 41% in the UC. At the same time, the number of letter writers using *ue* at least once rises from none at all in the LC to 47% in the UC, which is also the only social rank where more letter writers use *ue* at least once than use *ul* at least once. The distribution displayed in Figure 3 clearly corroborates the results presented in Figure 2, and we again conclude that *ue* entered, in social terms, from above.

As we have seen in Chapter 4, and will see in Chapters 7 and 8, changes from above were often also gendered, in that male writers from the upper and upper middle classes were more likely to use incoming variants or new conventions of letter writing. This also applies to the distribution of *ul* and *ue*. In Table 9, we present the distribution of the 1,606 epistolary forms (*ul* and *ue*) in letters written by men and by women, both in absolute and in relative numbers.

As can be seen in Table 9, there appears to be a strong gender effect: male letter writers using an epistolary form choose the new variant *ue* in one-third of all cases, whereas women adopting an epistolary form keep to *ul* in 91% of all cases. In order to be certain that no interference from social class is involved, the results for gender are divided into results by social class in Figure 4, where we opted for a different way of visualisation than previously because of the larger number of variables incorporated into the figure.
Chapter 6. Forms of address

Figure 3. Proportion of writers who use *ul* and *ue* at least once, by social class.

Table 9. The proportion of *ul* and *ue* within the category of epistolary forms of address, by gender.

<table>
<thead>
<tr>
<th></th>
<th>Male letter writers</th>
<th>Female letter writers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>ue</em></td>
<td><em>ul</em></td>
</tr>
<tr>
<td>N %</td>
<td>N %</td>
<td>N %</td>
</tr>
<tr>
<td>Total</td>
<td>384 36</td>
<td>686 64</td>
</tr>
</tbody>
</table>

Figure 4 demonstrates that the gender effect holds when the results are divided according to social class. The use of *ul* by women decreases when the social class index rises, while in the UC and UMC in particular the proportion of *ue* rises from approximately 10% to 20%. The lines representing male usage show a similar pattern in the LC and LMC. In the UMC, however, *ul* and *ue* approach each other, while UC men prefer *ue* in over 70% of all epistolary forms, diminishing the proportion of *ul* to less than 30%. UC men constitute the only group where *ue* outnumbers *ul*. In other words, the rise of *ue* implies, at least to a certain extent, a gendered change from above.

If *ue* is indeed an incoming variant replacing the older form *ul*, we are witnessing a change in progress and would also expect an age difference in the use of *ue* and *ul*, with relatively young writers showing a larger proportion of *ue* than relatively old writers. On the other hand, we have not established many convincing differences between age groups (see also Chapter 11), as gender and social class prove to be much more
Figure 4. The proportion of *ul* and *ue* within the category of epistolary forms of address, by gender and social class.

Table 10. The proportion of *ul* and *ue* within the category of epistolary forms of address, by gender and age.

<table>
<thead>
<tr>
<th>Age</th>
<th>Male letter writers</th>
<th></th>
<th></th>
<th>Female letter writers</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>ue</em></td>
<td><em>ul</em></td>
<td>Total</td>
<td><em>ue</em></td>
<td><em>ul</em></td>
<td>Total</td>
</tr>
<tr>
<td>&lt;30</td>
<td>131</td>
<td>30</td>
<td>313</td>
<td>70</td>
<td>444</td>
<td>18</td>
</tr>
<tr>
<td>30–50</td>
<td>196</td>
<td>42</td>
<td>267</td>
<td>58</td>
<td>463</td>
<td>14</td>
</tr>
<tr>
<td>&gt;50</td>
<td>22</td>
<td>22</td>
<td>76</td>
<td>78</td>
<td>98</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>349</td>
<td>35</td>
<td>656</td>
<td>65</td>
<td>1,005</td>
<td>50</td>
</tr>
</tbody>
</table>

reliable predictors of variation and change. In Table 10, we present the results for male and female writers according to the three age groups we distinguish: <30, 30–50, and >50. Note that the total numbers, viz. 1,005 forms for men and 536 forms for women do not add up to 1,606 epistolary forms, as we were not able to determine the age of every letter writer in the seventeenth-century autograph corpus.

No clear pattern emerges. All male age groups score about two-thirds to three-quarters of *ul*, which appears more or less in line with the overall proportion of 65% *ul* in letters by men (cf. 64% in Table 9 above). The fact that the 30–50 age group scores slightly less, viz. 58%, might be related to the fact that there are many middle-aged merchants in our corpus, who are usually assigned to UMC. As we have seen in
Section 3.1, business letters tend to contain more *ue* than private letters, and UMC men may have transported this tendency to their private letters. There is also some evidence that >50 men are more reluctant to adopt *ue* than younger men, but the numbers are not as evident as those pertaining to the variables of gender and social class. Turning to the female letter writers, a strong influence of age becomes even less probable. The scores for the two youngest age groups are in line with the general pattern for women, viz. over 90% *ul*. It is only in the >50 group that the score diverges, which difference is largely explained by the social class index of these female writers: almost all instances of *ue* by >50 women are produced by UC women. We conclude that the results do not provide us with convincing evidence for much influence of the variable age on the use of epistolary forms of address.

4.2.2 *The eighteenth century*

In the eighteenth century, 4,110 out of 5,692 terms of address were epistolary forms, amounting to 72% (cf. Table 7 in Section 3.4). In Table 11, we present the share of epistolary forms within the total number of forms of address according to social class.

In the LMC, the UMC and the UC, epistolary forms make up the majority of all forms of address. In the UMC and the UC in particular, their share convincingly outnumbers other forms of address such as *gij* and *u*. Moreover, the proportion of epistolary forms in the LMC, the UMC and the UC is higher here than it was in the seventeenth century (cf. Table 8 in Section 4.2.1). Surprisingly, the LC diverges from this pattern, as the proportion of epistolary forms drops from 53% in the seventeenth century to 34% in the eighteenth century. We will return to this in Sections 4.2.4 and 4.3. First, we will discuss the distribution of *ul* and *ue* in the eighteenth-century subcorpus.

At this instance, we should point out that the *ue* form encompasses a variety of options. Whereas the seventeenth-century data do not show any full forms such as *u edele* or *uwe edelheid* “your honour”, from which the abbreviation *ue* sprung (Section 2), the eighteenth-century subcorpus provides examples of full(er) forms, both as subject (13), as object including indirect object (14), and following a preposition (15).

<table>
<thead>
<tr>
<th>Table 11. Epistolary forms of address, by social class.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Epistolary forms of address</td>
</tr>
<tr>
<td>----------------------------</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>LC</td>
</tr>
<tr>
<td>LMC</td>
</tr>
<tr>
<td>UMC</td>
</tr>
<tr>
<td>UC</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>
(13) *als uwelEd er was*

if your very honour there were
‘if you were there’

(14) *ik hebbe UWelEdellen uit berbices Een gebrekkige brief geschreen*

I have your very honour from Berbice a faulty letter written
‘I have written a faulty letter to you from Berbice [Guyana]’

(15) *heeden ontfang ik Mijn Eerste Missivie van UWelEdelle*

today receive I my first letter from your very honour
‘today, I receive my first letter from you’

The eighteenth-century subcorpus contains 4,057 tokens of the *ue* type, 365 of which are fuller forms as in (13–15). This amounts to 9% of all epistolary forms. Considering the fact that no full forms were attested in the seventeenth-century subcorpus, as well as their fairly modest share in the eighteenth-century subcorpus, we decided to regard them as instantiations of the *ue* type. Importantly, the full(er) forms only occur in letters from the UC or UMC, with a vast majority of 309 out of the 365 tokens in UC letters. The use of these fuller forms was largely restricted to addressing persons with an academic degree or noble title, and these are mainly found in the UC.\(^\text{10}\)

As explained in Section 4.1, eighteenth-century usage shows less variation in the use of forms of address than seventeenth-century usage, and *ue* emerges as the preferred variant overall, replacing other forms such as *ul*. In Figure 5, we present the proportion of *ul* and *ue* within the total number of eighteenth-century epistolary forms, according to social class. The picture is clear: if an epistolary form is used, it is most likely *ue*, and this applies to every social class. The incoming variant has largely replaced the older variant. Occasional occurrences of *ul* are mainly found in the LC and the LMC. Figure 5 thus confirms our conclusions from Section 4.2.1, namely that the new variant *ue* entered from above. This is also corroborated by Figure 6, which shows the proportion of writers in each social class who use *ul* and *ue* at least once. In the UC, almost all letter writers use *ue* at least once, and none use *ul*. The proportion of writers that use *ue* at least once drops steadily in accordance with the social class index. At the same time, the proportion of writers who use *ul* at least once rises, reaching 5% in the LC. Note that about 70% of all LC writers still use *ue* at least once, despite the fact that the proportion of epistolary forms in the total results for forms of address is only 34% in the LC (cf. Table 11 above).

The eighteenth-century results so far corroborate our conclusion that *ue* was introduced from above. In Section 4.2.1, we also established gender differences, claiming that there was a strong gender effect on the distribution of *ul* and *ue*, in that *ue* did not

\(^{10}\) See Bax (2010) for a discussion of the use of similar titles in Dutch elite correspondence of the seventeenth century. An early exploratory study of seventeenth-century elite correspondence is Daan (1982).
Figure 5. The proportion of *ul* and *ue* within the category of epistolary forms of address, by social class.

Figure 6. Proportion of writers who use *ul* and *ue* at least once, by social class.
just enter from above, but was preferred by UC men in particular. Table 12 gives the eighteenth-century distribution of \(ue\) and \(ul\) for male and female letter writers, according to social class, and in both absolute and relative numbers.

As can be seen in Table 12, male letter writers of the LMC, the UMC and UC use \(ue\) almost every time they use an epistolary form. The social difference, shown in Figures 5 and 6 above, is revealed in the LC score, where the proportion of \(ue\) drops to 86%. In the letters written by women, the social division is found one layer higher on the class index. While UC and UMC women use \(ue\) virtually every time they use an epistolary form, the proportion of \(ue\) drops to 79% in the LMC. It is again 100% in the LC, but it should be recalled that the group of eighteenth-century LC women only contains two letters by two different writers, and the scarcity of sources may well distort our view. Looking at the different social divisions of male and female letter writers in Table 12, it seems rather likely that we have again found evidence of what we have called *incremental writing experience* (cf. Chapter 4, Section 4.3): women’s scores are often similar to men’s scores one rank down the social class index. In this case, the score of LMC women corresponds to that of LC men, and the score of UC and UMC women to that of LMC men as well as UC and UMC men.

As in Section 4.2.1, we also investigated possible age effects, assuming that a change in progress will relegate the older variant to the language of relatively older people. Since the number of \(ul\)-tokens, at only 35 instances, is rather small, we refrained from incorporating gender and/or social class into the following Table 13, and merely looked at the proportion of \(ul\) and \(ue\) in the age groups. The differences between the age groups are very small and far less obvious than the social and gender effects discussed above. We conclude, as we did for the seventeenth-century results (Section 4.2.1), that age is a less important variable in our sources than social class and gender. In the next section, we will correlate the seventeenth- and eighteenth-century results.
Table 13. The proportion of *ul* and *ue* within the category of epistolary forms of address, by age.

<table>
<thead>
<tr>
<th>Age</th>
<th>Male and female letter writers</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>ue</em></td>
<td>N</td>
</tr>
<tr>
<td>&lt;30</td>
<td>1,386</td>
<td>99</td>
</tr>
<tr>
<td>30–50</td>
<td>1,910</td>
<td>100</td>
</tr>
<tr>
<td>&gt;50</td>
<td>761</td>
<td>97</td>
</tr>
<tr>
<td>Total</td>
<td>4,057</td>
<td>99</td>
</tr>
</tbody>
</table>

4.2.3 Social effects on the change from *ul* to *ue*

As argued in Sections 4.2.1 and 4.2.2, *ue* came in from above, mainly through UC men, and spread from there to the language use of other social layers and of women, thereby replacing the older epistolary variant *ul* and, to a lesser extent, *gij* and *u*. Figure 7 shows the proportion of writers in each social class that use *ue* at least once, in both the seventeenth and the eighteenth century. This figure clearly demonstrates the spread of *ue* from the UC to the other social classes.

![Figure 7](image-url)
The change from *ue* to *ul* went hand in hand with an increase in the use of epistolary forms. Recall from Section 3.4 that the proportion of the epistolary forms *ul* and *ue*, relative to the total number of forms of address, rises from 58% in the seventeenth century to 72% in the eighteenth century, due to the sharp increase of *ue*. The increase of epistolary forms is shown in Figure 8. Figure 8 shows how the use of epistolary forms rises in the UC from c. 50% in the seventeenth century to more than 85% in the eighteenth century, in the UMC from slightly over 60% to 70%, and in the LMC from slightly less than 50% to over 55% (cf. also Tables 8 and 11 in Sections 4.2.1 and 4.2.2). In the LC, however, where 70% of the letter writers use *ue* at least once (Figure 7), the proportion of epistolary forms nevertheless drops from over 50% in the seventeenth century to just over 30% in the eighteenth century. To explain this, we have to look into the contexts in which epistolary forms appear (Section 4.2.4), as well as into the use of *gij* and *u* (Section 4.3).

### 4.2.4 The contexts in which epistolary forms appear

In the previous sections, we described diverging patterns, in that the LMC, UMC and UC show an increase in the use of epistolary forms, while the LC shows a decrease in the use of such forms. In order to understand these diverging patterns, we have to look into the contexts in which epistolary forms are used, and more in particular into the contexts in which LC writers use epistolary forms of address.
In the seventeenth-century autograph subcorpus, we found 80 epistolary forms in the LC, all of which are tokens of *ul* (Section 4.2.1). On closer inspection, it turns out that 52 of these tokens (65%) occur in epistolary formulae as described in Chapters 3 and 4. Examples include text-structural formulae (16), (subordinate) health formulae (17, 18), greeting formulae (19) and closing formulae (20).

(16) *Ick ul moeder leentien Cornelijis laat mijn bemijnde soon weten als dat*
I your mother Leentien Cornelijs let my beloved son know as that
‘I, your mother Leentien Cornelijs, let my beloved son know that’

(17) *als dat Ick tegen woordych noch met ul vader en suster en broeder In as that I at present still with your father and sister and brother in goede gesontheijt zijn good health am*  
‘that I, with your father and sister and brother, am still in good health at present’

(18) *als Ick oock hoopen dat het met ul oock soo Is as I also hope that it with you also so is*  
‘as I also hope that you are also like that [viz. healthy]’

(19) *groet ul susters en mijn moeder oock greet your sisters and my mother also*  
‘please send my regards to your sister and my mother as well’

(20) *hondert dusent goedenen nacht bij mij ul broeder Jochem Jansz hundred thousand good nights by me your brother Jochem Jansz*  
‘one hundred thousand good nights by me, your brother Jochem Jansz’

Epistolary forms are, however, also used in non-formulaic contexts in LC letters (21–22).

(21) *ick sou ul dijt goeet wel eer gesturt hebben I would you this stuff well earlier sent have*  
‘I would have sent you these things earlier’

(22) *daer was de ijonman op die mij ul brief mee bracht there was the young-man on who me your letter with brought*  
‘there was the young man on [that ship] who brought me your letter’

Formulaic expressions clearly attract epistolary forms of address, since other forms of address are only occasionally found in epistolary formulae. One rare example is (23).

(23) *Ick laet u weeten als dat I let you know as that*  
‘I let you know that’

For comparison, we also investigated the contexts in which *ul* and *ue* appear in seventeenth-century autographs from the UC. Out of 87 tokens of *ul*, 29 occur in an
epistolary formula (33%) and out of 59 tokens of *ue*, 17 appear in a formula (29%). In total, this means that 46 out of 146 epistolary forms of address (32%) occur in a formulaic context, which constitutes a clear difference from the 65% in LC letters. This may not be surprising, as epistolary formulae are less frequent in the UC than in the LC (see Chapter 4), but it still means that epistolary forms spread more pervasively to non-formulaic contexts in UC letters. This observation becomes particularly relevant when taking into account the eighteenth-century results.

In the eighteenth century, the LC picture resembles that of the seventeenth century: out of 111 epistolary forms of address in the LC (99 tokens of *ue* and 12 tokens of *ul*) 58 occur in a formulaic context, amounting to 52%. Thus, while the use of formulae decreases over time, also in the LC (Chapter 4), there continues to be a strong connection between epistolary forms of address and the formulaic context. This explains why the proportion of epistolary forms decreases in the LC with the diminishing use of formulae over time. As such, the language in the LC letters on the one hand, and the language in the LMC, UC and UMC letters on the other hand, evolve in opposite directions. In the UC, UMC and LMC, the decrease in the use of formulae is even stronger, but at the same time, the frequency of epistolary forms of address rises. In these social ranks, the epistolary forms of address spread to non-formulaic contexts to a much greater extent. Recall that in the eighteenth-century UC, no less than 86% of all forms of address were epistolary forms, viz. tokens of *ue* (Table 11, Section 4.2.2) and that, at the same time, in the UC, many formulae were only used sporadically or even not at all (Chapter 4). By way of illustration, consider (24), an excerpt from a letter by a female writer from the UC, who signs “A van Haeften”. In the first, fairly formulaic paragraph, she uses *Ued* and continues to use it in the second, non-formulaic paragraph, where it is only at the end that we find one token of *u*. Clearly, Van Haeften’s preferred variant is *Ued*, whether the discourse is formulaic or not.

(24) *met zeer veel aangenaamheid heb ik weederom het genoeghe* with very much pleasure have I again the pleasure
*gehad UEd zeer vriendelijke lettere van den 15 october 1779* had your very friendly letters of the 15 October 1779
*zoo orgéneel als dublikaat zeer wel ontfangen te hebben* so original as duplicate very well received to have
*waar uit ik met zeer veel blydschap UEd beider welstand* where from I with very much happiness your both well-being
*beneevens die van UEd twee lieve kindere vernoomen hebben* next that of your two sweet children understood have
*hebbe de Eer UEd met de Jongst geboorene van herte te* have the honour you with the youngest born from heart to
*fieliesiteere [...]* congratulate
*dog heb ik tevens met het uiterste leetweeze Uit UEd mesciéve* but have I also with the greatest sorrow from your letter
I hope that the Lord once for you provide will and on his time also once a good outcome will grant
the Lord wish I grant you there for patience and endurance to you on the way that the Lord with you wishes to keep to submit
‘with much delight, I have again had the pleasure of receiving well your very kind letters dated 15 October 1779, both the original and the duplicate, from which I have understood with very much happiness the well-being of both of you as well as that of your two sweet children. Have the honour of cordially congratulating you with the youngest [...] But I have also understood from your letter with the greatest sorrow that you are not really making progress at all, with which I hope the Lord will one day provide you, and grant you a good outcome in due course. I wish the Lord grant you patience and endurance to submit to the path the Lord wishes to offer you’

The developments described in the previous sections and in this section result in a social split, recalling the results in Chapter 4, Section 4.1. In the LC, there is a fairly strong connection between epistolary formulae and epistolary forms of address, to the extent that 65% of the seventeenth-century epistolary forms of address appear in formulae, and in the eighteenth century this number is still 52%. The difference is far from absolute, but epistolary forms are more likely to occur in formulaic contexts, and other terms of address are more likely to occur in non-formulaic contexts. In the other social classes, epistolary forms of address are much more likely to occur in non-formulaic contexts as well, and are thus generalised to any part of the letter. These divergent patterns are, obviously, linked to the distribution of *gij* and *u*, to which we will turn in the next section.

### 4.3 The pronominal forms of address

In this section, we discuss the distribution of the pronominal forms of address *gij* and *u*. Recall that *gij* is only used in the nominative, while *u* is its historical counterpart in the oblique and possessives. The form *u* has, however, invaded the nominative, as a result of which it is the present-day V-form in the standard variety. Obviously, the distribution of *gij* and *u* is not an isolated issue, but is closely connected to the distribution of the epistolary forms as discussed in the previous sections. In the preceding
sections, we focused on the epistolary forms since they make up the majority of all forms of address in both the seventeenth and the eighteenth century. Here, we will concentrate on the pronominal forms, interpreting their distribution against the background of the results presented so far.

Consider Table 14, where we repeat the proportion of epistolary forms relative to all forms of address in both centuries, and have added the proportions of \( gij \) and \( u \). As we have seen in the previous sections, the proportion of the epistolary forms increases in the LMC, UMC and UC, but not in the LC, where it actually drops quite impressively from 53% in the seventeenth century to 34% in the eighteenth century. As to \( gij \), there appears to be a social divide in the seventeenth century (see Nobels 2013: 92), with the two lower ranks using \( gij \) more frequently (22% and 23%) than the two upper ranks (about 15%). In the eighteenth century, numbers for \( gij \) drop in every social class, while the social pattern stabilises in a neat distribution, with gradually fewer tokens higher up the social scale. The same neat distribution can be seen in the eighteenth-century results for \( u \). The seventeenth-century results for \( u \) are, however, less obvious, and call for discussion, to which will we turn below. What should be noted here with regard to the diachronic development is that the proportion of \( u \) drops severely in the UC, stabilises in the UMC, increases slightly in the LMC, and rises sharply in the LC. Finally, the overall decrease in the use of both \( gij \) and \( u \) is directly connected to the increase in epistolary forms, as becomes evident from the total scores.

Turning to the results by social class, we see clearly that the UC generalises \( ue \) to all contexts, as argued in the previous sections. No less that 86% of all terms of address are epistolary, and are in fact \( ue \) in the UC. The relatively high score for \( u \) in the seventeenth century diminishes, giving way to \( ue \). In the UMC, a solid majority of 70% of all forms is epistolary and we notice a decrease in \( gij \), while \( u \) stabilises. The LMC is still less epistolary, and again \( gij \) decreases, but now \( u \) increases. In the LC, finally, \( gij \) also decreases, but so do the epistolary forms, while \( u \) rises to no less than 47%. In other words, whereas the other social classes prefer epistolary forms, and even more so than in the seventeenth century, the LC switches to \( u \).

<table>
<thead>
<tr>
<th>% Epistolary forms</th>
<th>% ( gij )</th>
<th>% ( u )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>17th c.</td>
<td>18th c.</td>
</tr>
<tr>
<td>LC</td>
<td>53</td>
<td>34</td>
</tr>
<tr>
<td>LMC</td>
<td>48</td>
<td>56</td>
</tr>
<tr>
<td>UMC</td>
<td>64</td>
<td>70</td>
</tr>
<tr>
<td>UC</td>
<td>50</td>
<td>86</td>
</tr>
<tr>
<td>Total</td>
<td>58</td>
<td>72</td>
</tr>
</tbody>
</table>
Importantly, the LC links up with the other social ranks with regard to subject positions, as suggested by the decrease of $gij$. In the seventeenth-century LC letters, we find 40 nominative terms of address, 35 of which (88%) are tokens of $gij$. The other 5 are tokens of $ul$. In the eighteenth-century LC letters, however, $gij$ is used in 57 out of 91 nominative forms of address (63%), while the epistolary forms take up 24 tokens (26%), and the other 10 being marginal forms such as $jj$. This shows that, in the LC, epistolary forms actually increase in the nominative at the expense of $gij$. In this respect, the LC results are similar to the results for the other social classes. In oblique and possessives, however, the LC shows a strong preference for $u$ in the eighteenth century, while all other ranks are moving towards the epistolary forms.

Summing up: the UC moves towards $ue$ in all contexts, relegating $gij$ and $u$. The same applies to the UMC, especially in nominatives, where $gij$ is pushed back, but $u$ stabilises in oblique/possessive functions. The LMC reduces the use of $gij$ in favour of epistolary forms too, but in oblique or possessives $u$ rises. This also applies to the LC, where the rise of $u$ is even stronger. What does this mean?

Since $gij$ and $u$ seem to behave differently, with $u$ in particular behaving quite unexpectedly, we first carried out a small case study of forms of address used in possessives. We began with the eighteenth-century results, where the distribution of $u$ seems to be least problematic, and in any case more or less in line with the distribution of $gij$, following the same social pattern of a decreasing proportion of $gij$ and $u$ when the social class index rises (cf. Table 14). Table 15 presents the eighteenth-century distribution of possessive forms of address across social class, focusing on the two main variants, viz. the pronominal form $u$ and the epistolary forms $ue$ and $ul$.

As can be seen in Table 15, there is a preference for epistolary forms in possessive function in the LMC, UMC and UC, completely in line with the overall results for these social ranks. The LC, however, displays the opposite pattern, and shows a strong preference for $u$ instead. This is one context where the decrease of epistolary forms in the LC becomes obvious. To gain a better understanding of what is going on here, we decided to compare the two extremes, viz. the scores for LC and UC, to the seventeenth-century results for possessives; see Table 16.

<table>
<thead>
<tr>
<th></th>
<th>u</th>
<th>%</th>
<th>ue/ul</th>
<th>%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>LC</td>
<td>72</td>
<td>70</td>
<td>29/2</td>
<td>30</td>
<td>103</td>
</tr>
<tr>
<td>LMC</td>
<td>120</td>
<td>41</td>
<td>165/11</td>
<td>59</td>
<td>296</td>
</tr>
<tr>
<td>UMC</td>
<td>110</td>
<td>24</td>
<td>343/5</td>
<td>76</td>
<td>458</td>
</tr>
<tr>
<td>UC</td>
<td>121</td>
<td>16</td>
<td>657/0</td>
<td>84</td>
<td>778</td>
</tr>
</tbody>
</table>

Table 15. Possessives in the eighteenth century.
Table 16. Possessives in the seventeenth century.

<table>
<thead>
<tr>
<th></th>
<th>\u</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>LC</td>
<td>9</td>
<td>23</td>
<td></td>
</tr>
<tr>
<td>UC</td>
<td>64</td>
<td>66</td>
<td></td>
</tr>
</tbody>
</table>

Table 17. Oblique forms of address in the LC in the seventeenth and eighteenth centuries.

<table>
<thead>
<tr>
<th></th>
<th>\u</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>LC – 17th c.</td>
<td>10</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>LC – 18th c.</td>
<td>75</td>
<td>58</td>
<td></td>
</tr>
</tbody>
</table>

As Table 16 shows, when compared to Table 15, the UC switches from predominantly \(u\) in the seventeenth century (66%) to predominantly epistolary forms in the eighteenth century (84%) – but this we already knew. This is the general pattern of the increasing use of epistolary forms. What is striking is the opposite switch in the LC, where a strong preference for \(ul\) in the seventeenth century (78%, see Table 16) becomes an equally strong preference for \(u\) in the eighteenth century (70%, see Table 15). In Table 17, we show that the LC makes the same switch in the case of forms of address in oblique functions. Here, the seventeenth-century preference for \(ul\) (76%) also becomes an eighteenth-century preference for \(u\) (58%). Again we may wonder how these seemingly divergent patterns should be interpreted.

We seek the solution in the general change from \(ul\) to \(ue\), described in Section 4.2. In this change, most social ranks switch to \(ue\), and the UC in particular appears to adopt a \(ue\)-only policy. To a lesser extent, this also applies to the UMC and the LMC. The LC, however, that reduces \(gij\) in the nominative in favour of epistolary forms, does not switch to \(ue\) in oblique and possessive functions, but to \(u\). It can therefore be seen that the change attested in LC language use is actually a change from \(ul\) to \(u\) in these contexts. Since \(gij\) and \(u\) may have been widely used in the spoken language, probably more so than the epistolary forms, and since they are in any case part of the historical second person paradigm, one could argue that LC language use develops from more written-like to more oral. But this would clearly be at odds with the changes in the nominative, where \(gij\) is dropped in favour of epistolary forms. It should also be recalled that no less than 70% of all LC writers in the presumably more “oral” eighteenth century still used \(ue\) at least once. Finally, our results in Chapters 2, 7 and 8 by no means suggest a development towards more “oral” language in the eighteenth century, but rather the contrary. Instead of assuming an evolution towards the spoken language, we would like to argue
that the change from *ul* to *u* should be interpreted as an intermediate stage in the change from *ul* to *ue*. What we are witnessing, is a reluctance to switch from *ul*, which is increasingly perceived as old-fashioned, straight to the incoming variant *ue*. The reason for this must be that *ue*, as the new form, is perceived, not as conventional or politic, but as polite or over-polite. This politeness or over-politeness is avoided by adopting *u*, a familiar and conventional form, and not as old-fashioned as *ul*.

In this interpretation, the puzzling distribution of *u* in the seventeenth century becomes crystal clear (cf. Table 14 above). Recall that the LC, LMC and UMC used *u* in about 20–25% of all cases, whereas the UC scored strikingly high with 34%. And on the other hand, the UC proportion of epistolary forms (50%) lagged somewhat behind relative to their proportion in the language of the UMC (64%). The same explanation holds here: in the change from *ul* to *ue*, which we know has taken place, there is an intermediate stage in which language users already drop *ul*, but are reluctant to use *ue*, and therefore choose *u* instead. Hence, the high frequencies of *u* for the UC in the seventeenth century, and for the LC in the eighteenth century. This intermediate stage can be explained by assuming that *ue* was perceived to be somewhat out of place in private correspondence, and the reader will have noticed that at this stage politeness strategies may play a role, after all.

5. Politeness – again

In Section 4, we have discussed, amongst other issues, a major change characterising the language of Dutch private letters of the seventeenth and eighteenth centuries, viz. the change from one epistolary term of address (*ul*) to another (*ue*). A second change is the overall increase in epistolary forms at the expense of the pronominal forms *gij* and *u*. Both changes are gendered changes from above, in that *ue* replaces *ul* first in the letters of UC men, spreading from there to the language of women and to other social ranks. In our seventeenth-century data, the change to *ue* is just taking off in UC language. In our eighteenth-century data, most social ranks have adopted *ue* in most contexts, but now the change is just taking off in the LC. In the seventeenth-century results for the UC and the eighteenth-century results for the LC, we see a similar, and at first sight surprising, effect of this change: the frequency of *u* is high, especially when compared to the *u* scores of other social layers. Therefore, we have argued that the change from *ul* to *ue* passes through an intermediate stage of *u*, so that, at least at the level of the population as a whole and not necessarily for every individual writer, the change should actually be conceived of as *ul* > *u* > *ue*. To explain this, we use the concept of politeness, and also draw on the results presented in Section 3.

In Section 3, we argued that politeness strategies cannot sufficiently explain the distribution of forms in our letter corpora. One central problem, for instance, is the abundance of supposedly negatively polite forms such as *ul* and *ue* in these private letters, many of which are exchanged between close family members. Nevertheless, we
did find relevant differences in the use of forms of address between specific groups of language users. One such clear difference existed between seventeenth-century private and business letters written by males from the UMC. Whereas private letters contained 67% epistolary forms, this proportion rose to 89% in business letters, an increase which was due to both *ul* and *ue*, but the rise of *ue* was even stronger than that of *ul* (cf. Table 2 in Section 3.1).

Towards the end of the seventeenth century, *ul* was a conventional, politic term of address in private letters. It may well have been an overtly polite form in the past, but in our view its abundance in our letter corpora can only be explained by assuming its straightforward conventionality. A new variant, *ue* came into use which at first may very well have been perceived as a polite or over-polite form. Furthermore, *ul* having become a neutral politic form, the need for a new polite or over-polite form is perfectly understandable. So when *ue* began its rise, it still had a polite or over-polite connotation, at least for some language users. These language users, willing to abandon the old-fashioned and/or merely conventional form *ul*, and conscious of the existence of a new variant, nevertheless did not switch to *ue* immediately, but chose *u* instead. This form, the historical oblique/possessive form of *gij*, was probably the politic form in the spoken language (cf. Section 2), and it therefore must have been a fairly safe choice in the language of letters. As *ue* spread, however, replacing *ul* in ever more contexts and situations, it became a conventionalised form itself. It lost, through what we might term pragmatic bleaching, its polite or over-polite connotation, and developed into the politic form. Towards the end of the eighteenth century, this was the situation for the UC, and perhaps for the UMC and LMC as well. In the LC, *ul* was perceived as out of place, as it was anywhere, but switching to *ue* was not an option yet for every LC writer, and again the frequency of *u* rose. In other words, for those LC writers, *ue* still had its touch of politeness or over-politeness.

Finally, *ue* was also being replaced, or at least this evolution had begun towards the end of the eighteenth century. Having lost its polite or over-polite significance, and being turned into a conventional, politic form, the need for polite or over-polite forms again arose. It should be recalled that there were no full forms of *ue*, such as *u edele*, in the seventeenth century subcorpus, but that we did find full forms in the eighteenth-century subcorpus. Out of 4,057 eighteenth-century *ue* tokens, 365 were full forms (9%, cf. Section 4.2.2) and among these 309 were found in letters from the UC. Thus, it may very well be the case that the UC was adopting new polite or over-polite forms.¹¹

¹¹ Ruberg (2005), discussing nineteenth-century usage, suggests that the full forms did not persist in the long run, as, quite the contrary, epistolary forms in general disappeared from the language of letters in the course of the nineteenth century. It is tempting to consider this as proof of the disappearance of *gij* and *u* from the spoken language in most situations. The *j*-forms may have become politic alternatives in everyday life, relegating *gij* and *u* to the written language, where *gij* and *u* were consequently transformed into epistolary forms of address, rendering *ue* and its full alternatives superfluous. A collection of nineteenth-century letters might shed more light on this.
6. Conclusions

In the present chapter, we have shown that investigating the forms of address used in the letters in our corpora from the angle of politeness strategies may shed some light on the distribution of the different forms. But we have also argued that from such a perspective it is difficult to fully explain the variation found. In particular, no satisfactory explanation was found to clarify why supposedly negatively polite forms abound in both centuries – in *private* correspondence. Taking a sociolinguistic perspective, we have revealed a clear distribution, patterning across social classes, and convincing diachronic developments. We saw epistolary forms increasing against pronominal forms such as *gij* and *j*-forms in the seventeenth century and even more strongly in the eighteenth century. The rise of *ue*, which began to replace the older form *ul* in the seventeenth-century corpus, was shown to be a gendered change from above, in that it was introduced by the upper ranks and spread from there to the lower ranks, and in that it was mostly men who led the change. For some language users, the change from *ul* to *ue* went through an intermediate stage of *u*. This development was interpreted as an effect of politeness strategies. At its incipient stage, *ue* must have had an overtly polite tone for these language users who therefore avoided it, especially in oblique and possessive uses, and preferred *u* instead. Interestingly, this initial reluctance to switch to *ue* was observed in different social classes. In the seventeenth-century data, the UC was switching to *ue*, and concomitantly its score for *u* was fairly high. In the eighteenth-century data, the same pattern was established in LC language use.
Chapter 7

Clause chaining between spoken and written language

1. Introduction

An intriguing aspect of the seventeenth- and eighteenth-century letters studied in the present book is the chaining of clauses by the use of semantically bleached connectives. Consider example (1), which is the beginning of the middle part of a seventeenth-century letter. Note that the transition from the preceding formulaic part to the middle part is marked by the concise text structural formula uoort soo “lit. further so”, discussed in Chapters 3 and 4.

(1) uoort soo laet ijc kl weeten dat ijc ende adam omtrent een maent further so let I you know that I and Adam about a month ijnde kooren bou gewrocht hebbe ijc hebbe nu een maent in-the corn cultivation worked have I have now a month voor korrelaer op snechtendeel geplocht ende hebbe noch wel een for Korrelaer on Schenectady ploughed and have yet well a maent werck te plochgen wij hebbe al omtrent twalf morgen taru month work to plough we have already about twelve morgen wheat geseijt ende ijc moet noch omtrent ses morgen met taru beseyen ende sown and I must still about six morgen with wheat sow and adam heeft nu drie weeken bij een ander boer de ploch gedreuen Adam has now three weeks at another farmer the plough driven ende moet daer noch uijf weeken blijuen dan sal hij verdijent hebbe and must there still five weeks stay then he will have veertijen schepel taru ende ijc uerdien omtrent smaents sestien fourteen bushel wheat and I earn about month+gen sixteen schepel taru ende als ijc gedaen hebbe met ploggen dan sal bushel wheat and if I done have with ploughing then will ijc sijen of ijc beg beginne te gorten aen het fort orangen I see if I begin to pearl barley+verb at the Fort Orange

‘Further, I let you know that I and Adam have worked in corn cultivation for about a month. I have now ploughed for Korrelaer in Schenectady for a month and I have still another month of work ploughing. We have already
sown about twelve morgens of wheat and I have still about six morgens of wheat to sow and Adam has been putting his hand to the plough at another farmer’s for three weeks now and he has to stay there for another five weeks. Then he will have earned fourteen bushels of wheat and I earn about sixteen bushels of wheat a month and if I am ready with ploughing, then I will see if I can start making pearl barley at Fort Orange’

The boldface instances of andaand” in this passage are examples of semantically bleached connectives chaining individual clauses. The ands do not coordinate units that need coordination in order to render the piece of discourse comprehensible. Each single andaand in boldface could have been omitted. Note, for instance, that the writer’s second statement, viz. that he has been working in Schenectady for a month, is attached to the first statement about his own and Adam’s recent activities in corn cultivation without any overt discourse marking. Separate statements are in some cases chained with andaand, and in other cases they are not. Thus, while andaand clearly coordinates discourse chunks in an expression such as ijck andaand Adam “I and Adam”, its main function in many other cases is rather to ensure discourse continuity by syntactic coordination. As Schiffrin (1987: 128) notes: “And has two roles in talk: it coordinates idea units and it continues a speaker’s action”.

This type of clause chaining with andaand recalls the well-known studies by Chafe. For instance, Chafe (1985: 111), discussing linguistic features from a corpus of spoken English, remarks the following:

In spoken language idea units are typically strung together in a chain, with a relatively small amount of subordination [...] Idea units may be independent [...] Or they may be linked by coordinating conjunctions, by far the most frequent of which is anda:

And it’s very well equipped.
You know the kitchen,
and and it’s got a dishwasher,
and it’s got all kinds of you know mixers and plates
and you know every kind of equipment you need.
And and staple things.

Chafe (1988) studies linkages between intonation units in present-day spoken English. In his corpus of spoken English, about 44% of the cases of unit linkages examined lacked overt verbal linkage, though not prosodic linkage such as intonation patterns, and in 56% of the cases “some more specific linkage marker is present [...]. Fully 50% of the cases with explicit connectives consist of or include the maximally general connective anda” (1988: 10). Stressing the continuative power of anda, Chafe proceeds: “The function of anda is evidently to signal only that the idea
expressed in the upcoming intonation unit moves forward from the idea that pre-
ceded” (1988: 11).  

We will return to idea units and similar notions in Section 6. What is important
here is that, with regard to historical written Dutch, we make the very same observa-
tion about chained pieces of discourse as Culpeper & Kytö (2010) make with regard to
Early Modern English. Culpeper & Kytö (2010: 158–183) study the use of clause chain-
ing and in Early Modern English from the perspective of genre variation, and conclude
that its frequency is relatively high in speech-based genres such as trials, drama and
witness depositions. It has been noted frequently that pragmatic markers in general are
“predominantly a feature of oral rather than of written discourse” (Brinton 1996: 33).
In the present chapter, we tie in with the approach taken by Chafe with regard to mod-
ern English, and by Culpeper and Kytö with regard to Early Modern English.

Clauses being linked by a semantically bleached connective such as and is a “com-
mon feature of narrative discourse” (Beaman 1984: 59), established for oral genres by,
among many others, Chafe (1985, 1988) and Beaman (1984). As has been said, it also
occurs in our letter corpora as shown in (1). It seems to be a prototypical case of oral
residue in a written text type (Ong 2002 [1982]: 36). In this chapter, we will examine
this phenomenon of clause linking with semantically bleached connectives and focus
on the use of ende and en “and”, which are also by far the most frequent empty connec-
tives in our letter corpora. We will determine the extent to which clauses are chained
by this relatively oral use of ende and en “and” in order to establish another degree of
orality. In Chapter 2, we focused on the degree of orality of seventeenth- and eight-
teenth-century private letters with regard to phonology, morphology and orthography.
In the present chapter, we will focus on clause linking.

Another way of achieving clause linking is by making use of punctuation. It is in
the period under discussion in the present study that punctuation increases in manus-
script spelling. This topic will also be touched upon in the present chapter. Greidanus

1. Beaman (1984) criticises Chafe, claiming that and should not always be considered a coor-
dinator. Consequently, spoken language might not be more coordinated than written language
as is often claimed. See Auer (2002) for a critique of this idea, and see also Chapter 8. Beaman
(1984: 59) gives the following example: And then he gets down out of the tree,/and he dumps all
his pears into the basket,/and the basket's full,/and one of the pears drops to the floor,/and he picks
it up,/and he takes his kerchief off,/and he wipes it out,/and places it in the basket/which is very full.
Beaman's comment runs that it ‘is clear that the ands in this example, particularly those initiat-
ing clauses, are serving as filler words or weak connectives parallelling the use of the adverb
then’. This is, however, exactly the kind of use of and we are interested in. The semantics of the
connectives is bleached, and its main function is continuing discourse.

2. About the different forms, we wish to point out that in Middle Dutch, the form ende was
commonly in use, recalling parallel Middle English and Middle Low and High German forms
such as ande and unde. Ende was gradually replaced by the reduced form en, but remained in use
well into the seventeenth century. In our seventeenth-century data, ende is the dominant form,
while en prevails in the eighteenth-century data.
(1926: 146–148), in her study of the history of punctuation in Dutch, claims that Middle Dutch clauses were often linked by *koppelwoordjes* “linking words” such as *ende* “and”, *so* “so” and *doe* “then”. These signalled the transition from one clause to another in the absence of any or at least any fully-fledged punctuation. Culpeper & Kytö (2010: 168) also point out that in the history of English the reduction of clause chaining *and* might be linked to the concomitant rise of punctuation as a visual means of clause combining and clause separation. The crucial period in this development is assumed to have been the mid-to-late seventeenth century.

In this respect, note that when we speak of clause chaining and clause linking as the main function of instances of *ende* and *en*, we also imply the separation of clauses. On the one hand, *ende* and *en* continue discourse and signal that the new clause proceeds from the preceding clause. On the other hand, they mark the transition from one clause to another, thereby separating them from each other. The same applies to punctuation marks such as commas and full stops. These may signal the end of a piece of discourse, but they also indicate the beginning of another clause, while simultaneously separating the ending from the upcoming clause.

In Section 2, the subcorpora used for the present chapter will be introduced. In Section 3, we will discuss the four strategies of linking clauses employed by letter writers in our corpus. These four strategies depend on the absence or presence of two clause chaining means, viz. semantically bleached connectives and punctuation. In Section 4, we will discuss the grammaticalisation of adverbs and conjunctions into semantically bleached clause linkages, and show that such grammaticalised usages occur frequently in our corpus. Note that the highly frequent forms *voorders*, *voorts*, *verder* and so on, meaning “further”, which could very well be considered bleached chaining connectives, were discussed in Chapter 4. In Section 5, we will discuss the rise of punctuation in manuscripts as another means of clause chaining and show that the rise of punctuation was probably a gendered change from above. Section 6 focuses on the use of *ende* and *en* “and” as a clause linkage, demonstrating that its use depends on variables such as social class, gender and time. In Section 7, we will discuss to what extent the two major developments in clause chaining, viz. the rise of punctuation and the decrease in use of *ende/en*, are interconnected.

2. The subcorpora

We compiled two subcorpora for the present case studies based on the seventeenth and eighteenth-century letters. We used the seventeenth-century autograph subcorpus (cf. Chapter 1, Section 6 and Chapter 4, Section 3.1), while limiting the number of letters by men from the upper-middle class (UMC) to the number of letters by women from the UMC, viz. 24. It should be noted that the seventeenth-century autograph subcorpus comprises 121 letters by UMC men. The restriction was necessary because of the laborious type of analysis used for the present chapter (see Section 6).
seventeenth-century subcorpus used here consists of 120 letters written by 107 different writers. Table 1 gives the structure of this subcorpus, divided horizontally into four social classes, and vertically into male and female writers. The social class index labels are lower class (LC), lower-middle class (LMC), upper-middle class (UMC), and upper class (UC). The number of letters (N Le) and the number of words (N Wo) are given for each social class and for each gender.

Table 2 gives the details of the eighteenth-century subcorpus used here, which is a subset of the total eighteenth-century corpus (cf. Chapter 1, Section 6 and Chapter 4, Section 3.1). For the UC and UMC men and women and for the LMC men, the number of letters was limited to the 24 letters by LC men. Note that the number of letters by LC women letters is low, even lower than in the seventeenth century. Fortunately, in the eighteenth century, men from the LC are far better represented than in the seventeenth century, which results in a reasonable total number of eighteenth-century LC letters. The eighteenth-century subcorpus consists of 160 letters written by 127 different writers.

3. Clause linking in the letter corpus

Chafe (1988) distinguishes between overt clause linkages through connectives such as and, and clauses linked only by prosody. Focussing on written language, we have to distinguish punctuation as a third possible means of clause linkage. The importance of punctuation as an alternative for verbal clause linkages is also noted by Culpeper & Kytö (2010: 168) in their study of and in various genres in the history of English.
Clauses may also be linked by semantically rich connectives such as *maer* “but”, *want* “for” and other Dutch coordinators, subordinators and adverbs which fulfil textual functions. These connectives will not be taken into consideration in the present chapter, in which we focus on grammaticalised connectives used for clause chaining, and particularly on *ende* and *en* “and”. This implies that in our letter corpora, we find four strategies of clause linking, depending on the presence or absence of semantically bleached connectives and of punctuation:

- no bleached connectives and no punctuation
- bleached connectives, but no punctuation
- bleached connectives as well as punctuation
- punctuation, but no bleached connectives.

A strategy closely related to punctuation is capitalisation of the first word of the sentence. This is yet another strategy that writers may employ. In fact, many though not all of the letters in our corpus seem to lack sentence-initial capitals, which suggests that in this period capitalisation as a strategy was just developing in manuscript spelling, as was punctuation. It is, however, notoriously difficult to distinguish minuscules from majuscules in manuscript spelling, in particular when there is no other formal difference than the size of the letter. For this reason, we will not take capitalisation into account.

In the following we give examples of the four strategies mentioned. It is important to note that letter writers do not necessarily use only one of these four strategies. They may vary, and connect, for instance, clauses without any marking (strategy 1), and then use bleached connectives further on (strategy 2), et cetera.

### 3.1 No bleached connectives and no punctuation

This first strategy of using neither bleached connectives nor punctuation occurs frequently in our corpora, though rarely consistently throughout a letter. Most writers do, at least occasionally, insert semantically bleached connectives. The opening passage of the following seventeenth-century letter (2) provides a good example of the first strategy. We give the diplomatic transcription of the first few lines of the letter, with glosses and a fairly literal English translation where we added punctuation and capitalisation in order to point out plausible transitions from one clause to another. The picture of the original manuscript (Figure 1) shows that the writer used a margin on the left side and wrote towards the right side of the paper as far as possible, chaining clauses by placing them next to each other, without making use of grammaticalised, semantically bleached connectives nor punctuation and/or capitalisation.

(2) adij in serruname 1672 den 12 septenber
    A.D. in Suriname 1672 the 12 September
    weerde moeder en broeders salut ick hebbe
    dear mother and brothers salut I have
Figure 1. Fragment of a letter without bleached connectives and punctuation.

noch 2 brieven geschreven een met soetelin
still 2 letters written one with Soetelin
en noch een met pieter ijaspers maer die van
and another one with Pieter IJaspers but that of
soetelin hebbe weer gekregen quam een huer te
Soetelin have back received came an hour too
laet waer uijt ul sult konnen verstaen onsen toe
late where from you will can understand our si-
standt in serruname sedert mijn laeste schrijven
tuation in Suriname since my last writing
hebben wij beije seer sieck geweest conen
have we both very ill been can
malkander weijnich helpen ...
each other little help ...
‘AD in Suriname 1672, the 12th September
Dear mother and brothers, salut. I have written another two letters, one [I
sent] with Soetelin and another one with Pieter IJaspers, but the one I sent
with Soetelin I have got back. [It] came one hour late. From [this letter] you
would have been able to understand our situation in Suriname. Since my last
letter we have both been very ill. [We] can hardly help each other’
In many instances we find writers chaining clauses as in (2), without any overt bleached linkages, though very often bleached connectives are used, as we will show now.

3.2 Bleached connectives, but no punctuation

The second strategy, also in use in example (1) above, is demonstrated in the next excerpt from the seventeenth-century (3) in which the writer reports on events related to the contemporary wars between Dutch, French and English troops. The bleached instances of *ende* are in boldface. Note that the passage begins with the formula *vors* “further” (see Chapters 3 and 4).

(3) *vors* weet Ick niet veel

*Further* know I not much

to write than the sad war here on

water and on land the French lie for our

frontier places well with a 20 thousand

men and here is today news come

that the English well 30 of our Setubal vessels taken and our fleet is at sea

but have still not by each other been

and the ship the Maeijboom is by the

English taken and Jan van Haerlijm he

lies in Bergen and dares yet not home

come by fear of the English...

‘Further, I do not know much to write other than the sad war here on the sea and on land. The French are stationed in front of our border towns with at least 20,000 men. And today the news came that the English have taken at least 30 of our Setubal vessels. And our fleet is at sea, but the ships have not yet regrouped. And the ship the Maeijboom has been taken by the English. And Ian van Haerlijm, he is stationed at Bergen. And he does not dare to come home yet for fear of the English’
The writer first connects clauses without any overt linking, thus employing strategy 1. The clause on the French troops being stationed close to the Dutch border towns is immediately connected to the foregoing, without punctuation. In the third line, the interspace between *lant* “land”, the final word of the first clause, and *de franse* “the French”, the first words of the new clause, does not seem to be particularly wide, though it is similar to other interspaces. The width of the interspace is not used to signal the clause boundary either and thus there is an unmarked transition from one clause to another. Further on, however, the writer repeatedly uses *ende* “and” for clause chaining.

3.3 Bleached connectives as well as punctuation

The third strategy is illustrated in (4), a passage from an eighteenth-century letter. Though there are quite a few writers using both bleached connectives and punctuation as a means of clause linking, it is not often the case that these two means are used simultaneously.

(4) *weegens ’t Niuwe Tractaat, die zy is ingtreeden, en myn* because of the new treaty, that she has entered, and my *Heer Cortiesos die gepasseerde week is hier gekoomen heeft* Sir Cortiesos who last week has here arrived has ‘because of the new marriage settlement, which she has entered upon. *And* Mr. Cortiesos, who arrived here last week, has [...]’
The writer refers to his sister and her new marriage settlement. The relative clause beginning with die “which” is separated from the antecedent ‘t Nieuwe Tractaat “the new marriage settlement” by a comma as is the next clause, which is, however, a new main clause about someone else, viz. Mr. Cortiesos. This new main clause is also linked to the preceding discourse by use of en “and”. The writer thus uses both bleached connectives and, occasionally, punctuation.

3.4 Punctuation, but no bleached connectives

In the next example (5), taken from an eighteenth-century letter, the writer separates the different clauses by making use of commas, including virgules, and full stops, as can be seen in both the transcription and the picture.

(5) UwEd: aangenaame van den 6. feb:
Your pleasant of the 6 February
deezes Jaars hebbe ik /op Gale Zijnde/
this+gen years+gen have I on Gale being
wel ontvangen, edoch hebbe het ge-
well received, however have the plea-
luk niet gehad omme op de Zelve
sure not had to on the same
te beantwoorden, verzoeke weshalven
to answer request therefore
mij te willen excuereen.
me to want excuse
Wijders maake uwEds bekend, dat
Further make your honour known that
‘Y our pleasant [letter] of February 6 of this year I have received well, while being at Gale. However, I have not had the pleasure of answering that same [letter]. Therefore, I request that you excuse me. Furthermore, I announce to your honour that [...]’

There are not many passages in the letter corpora where only punctuation and possibly capitalisation are used for clause linking. Punctuation in itself is a rare phenomenon, as we will see in Section 5. First, we will discuss in more detail the use of overt verbal clause linkages.
4. Grammaticalisation of clause linkages and discourse markers

In the present chapter, our main focus is on the semantically bleached clause linking use of *ende* and *en* “and”. This use of *en* and *ende* results from a grammaticalisation pattern by which propositional content is lost, and which also affects other conjunctions and adverbs such as adversative *maar* “but”, causal *want* “for, since”, consecutive *zodat* “so that”, temporal *nu* “now” and so on (cf. Brinton 1996: 50–65). We will discuss this grammaticalisation pattern in the present section, with examples from the subcorpora described in Section 2.

Connectives such as *en* “and”, *maar* “but” and *want* “for, since” combine cohesion in the sense of Halliday & Hasan (1976) with cognitive coherence in that they explicitly link clauses (and thus create cohesion) which are also logically connected (and therefore coherent), either through addition or by adversativity or causality. When these connectives grammaticalise, they lose propositional content: *maar* and *want* no longer necessarily introduce adversative and causal clauses, and *en* does not necessarily coordinate clauses at the same discourse level. In their grammaticalised function, they merely indicate discourse continuity by creating cohesion between clauses, while their function of establishing coherence is lost. Therefore, they may be used in non-coherent contexts as well, while still creating cohesion. As a consequence, the absence of propositional content allows for more possible contexts. As noted by Hopper & Traugott (2003: 94–98), semantic bleaching does not always imply just loss of meaning, but in many instances also implies pragmatic enrichment. This is also the case with the connectives under discussion. As they no longer depend on coherence relations, they can link any two clauses and, in this process, discourse marking functions
as described by Schiffrin (1987) develop. The wide range of discourse marking functions notwithstanding (cf. Alonso-Almeida 2008), we consider discourse continuity as the overarching function of the semantically bleached connectives concerned.

The following excerpt (6) contains examples of the grammaticalised discourse marker *maer*, which are set in boldface.

(6) *Ick heb ul*
I have you

*Ick weenst wel dat maer Ich quam op dit vaerwaeter niet*
as dear as my own heart but I wished well that I again with you was I came on this waterway not

*easyl weer maer Ick hoop dat dien almachtichgen легко again but I hope that the almighty Godt ons leits man sal wesen*

God our leader will be

'I love you as much as my own heart. But I wish that I were with you again. I would not soon enter these waters again. But I hope that the almighty God will be our leader.'

The writer expresses his love for the addressee and continues by expressing his hope that they will see each other soon. These two clauses do not contradict each other. On the contrary, if a coherence relation were to be assigned, it would probably be consecutivity. Nevertheless, the clauses are linked by *maer*. The writer proceeds by claiming that he would not go sailing again if they were together, and then expresses his hope that God will lead him. Again, *maer* is used as clause linkage, though unambiguous coherence between the two clauses is absent, as is unambiguous adversativity. One might reconstruct the use of *maer* by hypothesising that the waters referred to are dangerous, and that being led by God will help the writer to overcome the danger in which he finds himself. One might also hypothesise that being led by God will lead him back to his loved one, which is not where he is when writing the letter, but where he very much wishes to be. These are reconstructable contradictions in the discourse which could license the use of adversative *maer*. But one could also argue, as we do, that the clearest function of *maer* is to ensure discourse continuity, for which also other connectives with a similar function could have been used such as additive *ende* or *en “and”, then* (as Beaman 1984 notes) or the bleached causal connective *want* “for”.

As a semantically fairly unspecific, and a mainly discourse continuing connective, *maer* may develop into a real discourse marker orienting the hearer/reader into a certain rhetorical direction. The main function of this *maer* is to introduce a new topic. Consider example (7), taken from the same letter as example (6).

(7) *men*
man
seidt eens betaelt het al maer nu van wegen die
says once pay it all but now about those
ships
‘one says: pay it all. But now about those ships [...]’

The writer talks about payment and then shifts to a different topic, viz. particular ships. The topic shift is established by maer, reinforced by nu “now”. This topic shifting power of maer is also documented in historical dictionaries of Dutch, where maer and maar are said to introduce new topics or new discourse parts in the absence of a contradiction (MNW s.v. maer, WNT s.v. maar). In (8), maer is used three times in a row as clause linkage (in boldface), the first instance being the common adversative coordinator, the two next instances being examples of the topic shifting connective.

(8) Ick soude wel schrijven hoe Lange maer
I would well write how long but
Ick en wetet nijet den tijd sal Leeren maer hijer is sucker genoe[gh]
I NEG know-it not the time will teach but here is sugar enough
maer aengaende van ons kargesoen
but concerning about our load
‘I would write how long but I do not know. Time will tell. But there is sugar enough here. But concerning our load [...]’

Whereas the second and third instances of maer are examples of grammaticalised, semantically bleached clause linkage, they do maintain at a pragmatic level some of the adversative substance of the adversative coordinator maer in that the reader is directed towards a new topic, different from the preceding one.

Similar observations can be made regarding want “for”. Example (9) contains bleached clause linking want. Note that this excerpt from the formulaic closing parts of a letter begins with the concise formula (ende) voort “(and) further”.

(9) ende voort weest gegroet van antien gers ul nijcht met haer hvijs
and further be greeted from Antien Gers your cousin with her house
gesijn want zijn zijn noch gesondt ende wensschen ul het best toe
family for are are still healthy and wish you the best to
‘And further, be greeted from Antien Gers, your cousin and her family. For they are still healthy and wish you the best’

The fact that these people are healthy surely enables them to greet others, but their health is not the cause of the greeting. This clause chaining use of want is also documented in historical dictionaries such as MNW s.v. want, where its meaning is paraphrased as en toen “and then”.

As maer may contain a pragmatic glimpse of its original adversativity, likewise want may exhibit a glimpse of its original causality. In such cases, want does not coordinate a
cause and an effect, but it indicates discourse continuity while informing the reader/hearer that what follows has some explanatory power. Consider example (10).

(10) *dar ijck seer blijde om ben dat*  
there I very happy about am that  
*ijeij weder om gesont waren mar de heer de*  
you again healthy were but the Lord he  
*slat en heij genest weder den gene de hij be-*  
strikes and he heals again the one who he  
*left want hij casteijt een ijegelijck de heij lijef-*  
wishes for he castigates one every whom he loves  
*heft*  
‘about which I am very happy, that you are healthy again. But the Lord he  
strikes and he heals again the one he wishes. Because he castigates everyone  
whom he loves’

The writer expresses his happiness about the addressee’s regained health. Here, the occurrence of *mar* “but” is an example of the grammaticalised clause linking use without any opposition or contradiction. A causal connective would have been equally suitable. The use of *want* in (10) is intriguing, as the subsequent clause does not provide a cause but rather a reformulation of the religious moral expressed in the previous clause. It again explains the addressee’s perception of health, illness and the role of God, which explanation is indicated by originally causal *want*.

When discussing the semantically bleached clause chaining function of *want*, the historical dictionary *MNW* (s.v. *want*) refers to *sodat* “so that” which in present-day Dutch typically introduces consecutive subclauses. A similar bleaching of the consecutive meaning of *sodat* is claimed with reference to the following thirteenth/fourteenth-century example (*MNW* s.v. *sodat*):

(11) *Vele jaren hebben si in dese vaert gedaen, so dat si in een eylant*  
Many years have they in this voyage done, so that they in an island  
*quamen*  
came  
‘Many years they spent on this voyage, so that they came in an island’

The arrival at some island is not the consequence of a voyage of many years, except when the island in question can only be reached by such a long voyage. Note, however, the indefinite article in the phrase *een eylant* “an island”, indicating that the discourse is not about a particular island which the travellers wanted to reach. Here *sodat* only signals the continuation of discourse. Similar usages of this clause chaining *sodat* can be found in the letters presently under discussion. This clause chaining function of *sodat* is, without any doubt, related to the well-known clause chaining use of *soo* in conditional constructions. This so-called expletive or resumptive *soo* is exemplified in (12).
Chapter 7. Clause chaining between spoken and written language

(12) *doch als ick hier wilde blijven:*  
but if I here wanted stay  
*en conde in vrede leven*  
and could in peace live  
*soude ick wel een goede plantagije coopen of maken*  
so would I well a good plantation buy or make  
‘But if I would want to stay here and could live in peace, *so/then* I would surely buy or create a good plantation’

The consequence of the condition is introduced by *sou*, which may also be omitted, as in present-day standard Dutch *dan* “then”, and in present-day standard English *then*. Now consider example (13).

(13) *voorders mijn beminde soon soe laedt ick v l weten als dat mester ijacop blocken soon oock doot is ende soe heb ick oock soe daedelyck verstaen*  
*Blocken son also dead is and so have I also so now understood dat mester ijacop seluer doot is soe dat daer ael veel gestueruen*  
*that Master IJacop self dead is so that there already many passed away*  

‘sijn’  
*Further, my beloved son, so I let you know that Master IJacop’s son is dead as well, and so I have just now also understood that Master IJacop is dead himself, so that many have passed away’

The excerpt begins with the concise formula *voorders* “further”, which we have discussed in Chapters 3 and 4, where we also pointed out that it may or may not be followed by expletive *so(o)* or *soe*. This *soe*, as in example (13), may immediately follow *voorders*, or it may occur later, for instance after the vocative as in the first line in (13). There are two more clause chaining usages of *soe* in (13), the first in combination with *ende*, the second *soe dat*. As to the latter, the fact that many have passed away is not a logical consequence of Master IJacop’s death. At most, it is at an epistemic level a conclusion which the writer draws on the basis of a range of individual observations, such as Master IJacop’s death. With this switch from the content to the epistemic level (cf. Sweetser 1990), *soe dat* takes up a pragmatic function, where it continues discourse and indicates the speaker’s perspective.

The temporal adverb *nu* “now” may also adopt a clause linking function. Consider (14), where the first instance of *nu* is semantically empty. As a discourse marker, it draws attention to a different or new perspective on the matter at hand: the writer is not in the position of sending money, nevertheless he hopes that the addressee will manage. As the second instance of *nu* shows, it may also be used as a topic shifter.

(14) *ick wenste wel dat gij*  
I wished well that you
I do wish that you would get 40 to 50 guilders from me, but I dare not send it over with someone else. Now, I hope you will nevertheless manage, for there are always good people who will help you a little, for I myself I hope to use it, please God, for chestnuts. Now I do not know anything special to write anymore, except [...]’

The connectives discussed so far such as *maer nu* in (7) and *ende soe* in (13) may also be combined. Other combinations are also possible. In (15), the writer uses *en oock soo* “and also so” as a topic shifter three times in a row.

(15)  
want ick tot noch toe gheen bescheijt en  
for I until now to no answer NEG  
*hebben en oock soo* kan ick niet verneemen of de scheepen  
have and also so can I not learn if the ships  
*van oorloegh haest sullen uijt gaen ofte niet want men*  
of war soon will out go or not for man  
*hoorter hier gansch niet af en oock soo schijnt het dat*  
hears-there here absolutely not about and also so appears it that  
de scheepen noch niet klaer en zijn want het haer aen  
the ships yet not ready NEG are for it them of  
*volck manckeert en oock soo leggen daer noch vijf*  
person lacks and also so lie there still five  
scheepen van oorlogh  
ships of war  
‘For until now, I have not yet received an answer, and also so, I cannot learn if the ships of war will soon leave or not, for one does not hear anything about it here, and also so, it appears that the ships are not yet ready, for there is a lack of people, and also so, there lie another five ships of war’
Apart from a shift of attention or a change of topic, the connectives under discussion may also signal foreground and background information, or distinguish topics from events within topics, as in example (16).

(16) **vorder soo** *laet ick ul weetten als dat ick*  
*Further so let I you know as that I*  
*in Spanie gevange ben geweest oom dat ick*  
in Spain captured have been because I  
*niet heb wille dienen in de aermade en al*  
not have want serve in the armada and  
*ik ewel soo heb ick moeten dienst neemen*  
nevertheless so have I must service take  
*soo ben ick ruijter geworden en daer ged[...]*  
so have I rider become and there served  
*hebbende een yaer soo ben ick weer gevange*  
having a year so am I again captured  
*gewoerden van de portegiesen in poortegal*  
become by the Portuguese in Portugal  
*en daer sijnde soo hebbe ick weer dienst*  
and there being so have I again service  
*moete nemen*  
must take  

‘**Further so** I let you know that I was captured in Spain because I did not want to serve in the armada. **And** nevertheless, **so** was I forced to take service, **so** have I become a rider. **And** having served there for a year, **so** I was captured again by the Portuguese in Portugal. **And** being there, **so** I was forced to take service again’

The writer begins by introducing the topic of his imprisonment in Spain by the use of *vorder soo* “further soo” and adds the reason, viz. that he did not want to serve in the armada. Three chronologically ordered topics follow which are foregrounded by the use of *en* “and”. First, that he has still served in the army for a year; second, that he was captured by the Portuguese; third, that he joined the army again. Chaining *en* first introduces the fact that he still had to serve in the army. The second instance marks the end of that period and introduces what came next, so it signals a temporal boundary in his biography. The third instance marks the transition to another place, viz. Portugal, and thus signals a spatial boundary in his biography. Each instance of *en* is followed by one or two clauses with the relevant information on the topic, introduced by *soo “so”*. This way of ordering elements in a discourse recalls Schiffrin’s study of *and*, in particular where she notices a similar hierarchy of *so* as topic linkage on a global level, and *and* as a linkage between events on a local level, that is, within topics (1987: 139). Note that in (16) it is the other way around: *en* functions on a global level, indicating the major
Letters as Loot

topics in the writer’s biography, while *soo* connects the clauses locally connected to the topic at hand.

The last example in this section is an entire letter from 1664 with different clause linkages, which also illustrate the global/local distinction discussed above. In this case, however, *en* is not the main topic shifting and foregrounding connective as in example (16), but it rather links the clauses within a topic. Topics are mainly, though not exclusively, introduced by *voorts*. Consider example (17).

(17) *Eersame seer beminde lieve moeder ende broeders*  
Honourable very beloved dear mother and brothers  
*Ick laet ul weten als dat Ick noch gesont ben Ick*  
I let you know as that I still healthy am I  
*hoop dat het met ul en mijn twee broeders mede*  
hope that it with you and my two brothers also  
*soo is waer het anders het soude mij van herten*  
so is were it different it would me from heart  
*leet sijn omte hooren. Voorts wat hier passeert*  
sorrow be to-to hear. Further what here passes  
is niet veel daer leggen niet veel schepen en daer en sijn geen  
is not much there lay not many ships and there *NEG* are no  
*vraghten En de karstanjers die deugen niet veel*  
loads and the chestnuts these be good not much  
*voort heb Ick ul briefien wel ontfangen En daeruyt*  
further have I your letter+DIM well received and therefrom  
*verstaen als dat onse Jacop noch gesont was maer*  
understood as that our Jacob still healthy was but  
*onse Cornelis die hadt de koors voorts mijn beminde mijn*  
our Cornelis he had the fever further my beloved  
*moeder laet het ul niet swaer vallen Ick heb hier*  
mother let it you not heavy fall I have here  
*gen sal met hooghlantsche wijne voort soo laet ick*  
no need I hope that my father a load+DIM get  
*ul weten als dat de karstanjaers niet veel deuge*  
you know as that the chestnuts not much are good  
*mijn vader sou wel wat na huijs toesture maer sij*  
my father would well something to home send but they  
*deuge niet verhoope dat mijn vader self wat*  
are good not hope that my father self something  
*breng sal voort soo seght gom Claeses maertjen*  
bring will further soo say gom? Claeses Maertjen
ende de andere meisjes en de knechten en mijn broeder
and the other girls and the boys and my brother
Jacop veel goede nachten en laat hij u gehoordsaam zijn in
Jacob many good nights and let he you obedient be in
Alle dingen dat recht en billick is. Gij schrijft oock
all things which just and fair is you write also
dat onze Cornelis heel drock heeft met de henjes Gij
that our Cornelis very busy has with the chicks You
schrijft oock dat zij alle leggen en ick heb verstaen dat Cor
write also that they all lay and I have understood that Cornelis
nelis taen[...] also thuijs is ende segh Garbrant siercksz dat

Taems already at home is and say Garbrant Siercksz that
sijn broeder harck siercksz behouden hier gecomen is ende
his brother Harck Siercksz safely here come is and
Taems Cornelisz buysman Is al tot bajoenen gecomen
Taems Cornelisz Buysman is already at Bayonne come
Voorts mijn beminde moeder Ick en weet ul geen meer
Further my beloved mother I neg know you not more
teschrijven sijt hertelijck van mij gegroet en Godtalmachtich
to write be heartily from me greeted and God almighty
In zijn genade bevolen. Ul dienstwillgen soon
in his mercy commended your dutiful son
Claes Jansz kuijper
Claes Jansz Kuijper

‘Honourable very beloved dear mother and brothers. I let you know that I am
still healthy. I hope that you and my two brothers are healthy as well. If it were
different, I would very much regret to hear that. Further, what happens here
is not much. There are no ships and there are no loads. And the chestnuts are
not much good. Further, I have received your letter well, and understood
from it that our Jacob was still healthy, but our Cornelis had fever. Further, my
beloved mother, do not let it make you feel bad. I have no need here. I hope
that my father will get a load of highland wine. Further so, I let you know that
the chestnuts are not much good. My father would have sent something home,
but they are no good. I hope that my father will bring something himself.
Further so, say many good nights to [gom?] Claeses Maertjen and the other
girls and little boys and my brother Jacob and let him be obedient to you in all
things which are just and fair. You also write that our Cornelis has been very
busy with the chicks. You also write that they all lay eggs. And I have under-
stood that Cornelis Taems is already at home. And say to Garbrant Siercksz
that his brother Harck Siercksz arrived here safely. And Taems Cornelisz
Buysman already arrived at Bayonne. Further, my beloved mother, I do not
know anything more to write to you. Be cordially greeted from me, and com-
mented to God almighty in his mercy. Your dutiful son, Claes Jansz Kuijper’
First, it should be noted that the formulaic beginning and closing are separated from the non-formulaic middle part by *voorts* “further”. This first instance of *voorts* marks the transition to the middle part. Within the following piece of information, *en* is used as a linkage, *voort* for the shift to another topic, and again *en* is used as a linkage within the topic at hand. Subsequently, *voorts*, *voort soo* and *voort soo* are used for new topics, the last containing another instance of *en*. The writer also switches twice to the phrase *Gij schrijft oock* “you also write” which functions as a clause linkage. The second instance is again followed by *en* as a linkage. Finally, *ende* “and” is used twice as a topic-shifting and discourse-continuing connective, after which *voorts* marks the transition to the formulaic closing. In summary, this fairly short letter provides an excellent example of the patchwork-like way in which writers may turn the individual clauses they want to communicate into a cohesive piece of discourse.

The overarching function of the connectives discussed in this section is to ensure discourse continuity. Topics and events are linked by semantically fairly bleached connectives, for which coordinators such as *en*, *maer* and *want*, and adverbs such as *nu* are recruited. These lose propositional content, and thus the ability to create coherence at a logical level, but they still create cohesion. At the same time, they adopt discourse structuring functions and may develop into discourse markers, distinguishing, for instance, foreground and background information, or global topics and local events within topics. Despite the wide variation in usage, of which only a glimpse was discussed in the present section, and despite the various pragmatic functions that may develop, we maintain that the joint characteristic of all these forms is their semantically bleached function as clause linkage. This is not only clear from the fact that almost all of the connectives discussed could equally have been omitted, without loss of propositional meaning. First and foremost, it becomes clear from the sociolinguistic restrictions on their use that appear to exist in that groups of speakers may use them to a greater or, indeed, to a lesser extent. This will be the topic of Section 6, following a discussion of punctuation as another way of linking and separating clauses.

5. **The rise of punctuation**

As in the history of English (Salmon 1999), the Early Modern period was important in the rise of punctuation in Dutch manuscript spelling as well as in printed books. Greidanus (1926) shows how punctuation systems developed in the course of the sixteenth century, in particular in normative publications on Dutch orthography and grammar. Höchli (1981) discusses similar developments in the history of German. Of special importance is Greidanus’ (1926) main claim that punctuation developed from predominantly rhythmic to mainly syntactic. Rhythmic punctuation indicates plausible pauses in spoken language, and thus maintains a fairly straightforward connection between spoken language and its written counterpart. Rhythmic punctuation was common in primarily oral societies, such as the Low Countries in the Middle Ages,
where written texts were frequently read aloud. It should be noted that punctuation is in fact lacking in many if not most Middle Dutch texts, where, however, verse length may be seen as a means of rhythmic punctuation, indicating natural pauses. The transition towards general literacy, and the textualisation of society as a whole, is a post-medieval development. Concomitantly, punctuation systems became predominantly syntactic, indicating boundaries between main clauses by full stops, and between main clauses and subclauses by commas. In other words, punctuation syntacticised.3 Similarly, Culpeper & Kytö (2010: 168–170) describe the rise of punctuation in seventeenth-century English in terms of a transition from the period to the sentence. In spoken language, prosodic features such as intonation help to identify periods as units of discourse. In written language, punctuation marks indicate the visually and syntactically defined sentence (cf. Lennard 1995).

We will use this idea about the syntacticisation of punctuation as the background to our case study of punctuation. However, we will not embark upon different usages and functions of commas and full stops and other punctuation marks in the letter corpora under discussion: that is, we will not try to disentangle rhythmic punctuation systems from syntactic punctuation systems. The main reason for this is that punctuation as such is rare in our letter corpora: some writers use punctuation, but most do not. Furthermore, inter-speaker variation is prevalent, as in English (Salmon 1999), and moreover, writers tend not to be very consistent, rendering intra-speaker variation almost equally widespread. Therefore, we will focus not on a possible change from rhythmic to syntactic punctuation, but on the rise of punctuation itself. In our approach, in which we concentrate on what happens when people stop talking and start writing, that is, in which we focus on the transition from spoken to written language, it is interesting to see if there are any sociolinguistically relevant patterns in the distribution of punctuation. If punctuation is a prototypically written means of clause linking, is the use of punctuation then socially distributed in accordance with the social distribution of writing experience? This is the topic of the present section.

In Chapter 4, we discovered clear distributional patterns in the use of epistolary formulae. Formulaic language was more in use in the seventeenth century than in the eighteenth century, more in the lower classes than in the upper classes, and more in letters by women than in letters by men. We interpreted these diachronic, social and gender effects in terms of writing experience. Formulae were shown to function as a safe option for less-experienced writers, which are relatively overrepresented among women, in the lower classes, and in the seventeenth century. If punctuation characterises the written code and, therefore, is also linked to writing experience, we would expect similar distributional patterns in the rise of punctuation. Our hypothesis would therefore be that punctuation occurred more frequently in the eighteenth century than in the seventeenth century, more frequently in the upper classes than in the lower classes, and more often in letters by men than in letters by women.

3. See also Chapter 8 on the notion of syntacticisation.
Examining the use of punctuation in the subcorpora presented in Section 2 above, we distinguish two types of letters: those with and those without punctuation. It should be noted that there are hardly any letters with consistent punctuation throughout the whole letter. Most writers who use punctuation, do so irregularly, or in any case irregularly compared to punctuation in contemporary printed books. The writer of example (5), for instance, uses a comma before the main clause beginning with *verzoeken* “(I) request” and a full stop at the end of that clause. Often, writers vary not in the choice of a particular punctuation mark, but in the presence of punctuation. Example (4) above contains two relative clauses introduced by *die* “which/who”; the first is preceded by a comma, while the second is not. In short, letters with punctuation may contain only a handful of punctuation marks, but may also have commas or full stops at the end of virtually all clauses.

Table 3 gives the presence of punctuation in the subcorpora used for the present chapter across social class and gender. The figures indicate the number of letters with and without punctuation. As becomes clear from Table 3, the use of punctuation is not at all common in seventeenth-century nor in eighteenth-century letters. In the LC and the LMC, for instance, hardly any writers use punctuation. Overall, punctuation is used in 27% of the seventeenth-century letters. In the eighteenth century, the proportion of letters with punctuation rises to 42%. In Figure 5, we plotted the proportion of users of punctuation across social class and gender, as well as diachronically.

As Figure 5 shows, there is a clear social distribution in the use of punctuation. Men in the seventeenth century, represented by the black columns, rise from no punctuation at all in the LC to almost 24% of the letters in LMC, about 60% in the UMC, and 73% in UC. In the eighteenth century (the dark grey columns), numbers are consistently higher in the LC, the LMC and the UCM, while UC-men score about 80%.

<table>
<thead>
<tr>
<th></th>
<th>17th century</th>
<th>18th century</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No punctuation</td>
<td>Punctuation</td>
</tr>
<tr>
<td>LC-M</td>
<td>5</td>
<td>–</td>
</tr>
<tr>
<td>LC-F</td>
<td>5</td>
<td>–</td>
</tr>
<tr>
<td>LMC-M</td>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>LMC-F</td>
<td>19</td>
<td>–</td>
</tr>
<tr>
<td>UMC-M</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>UMC-F</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>UC-M</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>UC-F</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Total N</td>
<td>88</td>
<td>32</td>
</tr>
<tr>
<td>Total %</td>
<td>73</td>
<td>27</td>
</tr>
</tbody>
</table>
Chapter 7. Clause chaining between spoken and written language

The rise in punctuation was clearly a change from above. There is a convincing gender effect as well. In the seventeenth century (the middle grey columns), hardly any women use punctuation, but some in the UMC and the UC do. In the eighteenth century, still only a handful of women from the LMC and the UMC use punctuation (the light grey columns), but the steep rise in the UC is spectacular: from 14% in the seventeenth century to over 60% in the eighteenth century. Letters by women show consistently less punctuation than letters by men. Thus the rise of punctuation was not just a change from above, it was led by UC men in particular, followed by other men, and by UC women in the eighteenth century. There is only one score not in line with this general trend, which is the 50% score for LC women in the eighteenth century. This score, however, is based upon an extremely low number of letters, viz. 2 (see Table 3), one of which contains punctuation while the other does not.

Summing up, when focusing on the presence of punctuation, the textualisation of society must have promoted the use of punctuation first and foremost among those who were most involved in the written culture. Put differently, we expected that the rise of the use of punctuation would show social and gender effects in that more men than women, and more upper-class than lower-class writers would use punctuation. This was borne out by our case study, which showed that the rise of punctuation was a gendered change from above. Finally, the use of punctuation increased over time.

6. Variation and change in clause chaining

In this section, we concentrate on one of the bleached clause chaining connectives discussed in Sections 1, 3 and 4, viz. the highly frequent connective meaning “and”,

Figure 5. Proportion of letters with punctuation in the seventeenth and eighteenth centuries, by social class, gender and time.
which is commonly rendered *ende* in the seventeenth century, and *en* from about the eighteenth century onward. Throughout the present section we will simply speak of *en*. Our main aim is to detect possible variational patterns in the use of chaining *en* related to social variables such as class and gender. Brinton notes that “[b]ecause of their frequency and oral nature, pragmatic markers are stylistically stigmatized and negatively evaluated, especially in written or formal discourse” (1996: 33). Keeping in mind the results of the previous section on punctuation, as well as the results of the case studies on formulaic language (Chapter 4), we hypothesise to find relatively more bleached connectives in letters from the lower classes than in letters from the upper classes, more in letters by women than in letters by men, and more in letters from the seventeenth century than in letters from the eighteenth century. In other words, bleached connectives, as more oral characteristics of language, are supposed to be less prominent in letters by writers who are more involved in the written culture.

The forms *ende* and *en* belong to the most frequent connectives, and in fact to the most frequent forms in the entire corpus. Four main functions of *en* may be distinguished: two coordinative functions, chaining *en* and formulaic *en* (cf. Culpeper & Kytö 2010: 160–164). In (18), *en* coordinates two verbal phrases; in (19) two noun phrases are coordinated. In both cases, *en* cannot be omitted. In (18), two predicates and two distinct pieces of information are coordinated; elision of the subject and the finite verb occurs, and the connective *en* fulfils its common coordinating role. In (19), *ende* coordinates two constituents which together make up the direct object of the finite verb *wensch* “wish”.

(18) **nu sal ik van deeze afstappen en tot de ander over gaan**
    now will I from this step down and to the other over go
    ‘I will now abandon this subject and proceed to another one’

(19) **ick wensch ul geluck ende saligheijt**
    I wish you happiness and salvation
    ‘I wish you happiness and salvation’

Example (20) contains what we call chaining *en*, the topic of this section. It creates cohesion of two clauses and ensures continuity of the discourse from one clause to the other, but it may equally have been omitted, as it does not contribute to the propositional content. Culpeper & Kytö (2010: 161) refer to this use as clause-level coordination.

(20) **t is alles wel ontfangen en volgens uw versoek neef heuck gesprooken**
    it is all well received and according to your request cousin Heuck spoken
    ‘it is all well-received, and in accordance with your request I have spoken to cousin Heuck’

Example (21) presents what we call formulaic *en*. The string *doet soo wel en seght* “be so good and tell” is a fixed formula which appears frequently in the subcorpora. It
could, of course, be argued that this is an example of chaining en, but the meaning of the first clause doet soo wel completely depends on the second clause. Doing well only manifests itself in the act of telling. In other words, this is not just a fixed formula, but also a hendiadys. We will regard examples such as these separately.

(21) doet soo wel en sees de schiper
    do+imp so well and say+imp the skipper
    ‘be so good and tell the skipper’

In order to disentangle these four functions from one another, and to be able to quantify the results of chaining en, we decided to split the subcorpora used for the present chapter into discourse units. In Section 6.1, we explain this division into discourse units and, subsequently, in 6.2, we discuss sociolinguistic patterns found in the use of chaining en.

6.1 Discourse units

To divide the letters into intonation units, that is, into “sequence[s] of words combined under a single, coherent intonation contour, usually preceded by a pause” (Chafe 1987: 22), is, clearly, not possible in the case of written language. We therefore instead adopt the term discourse unit (cf. Degand & Simon 2009) and building on Chafe’s work on intonation units, we consider these discourse units as sequences of words “verbalizing the speaker’s focus of consciousness at that moment” (Chafe 1994: 63). In other words, apart from the formal characteristics in spoken language, that is, their marked prosody, intonation units are also characterised by functional properties. Their main functional property is that they prototypically contain one new idea (Chafe 1994: 108–119), a claim which ties in, as Chafe (1987: 32) mentions, with Pawley & Syder’s (1983: 564) early suggestion that “[t]here may be a fundamental limit in humans’ verbal processing which may be termed the one-clause-at-a-time-constraint”. This constraint, in its turn, matches Chafe’s claim that many intonation units ‘have the grammatical form of single clauses’ (1994: 65). In the process of dividing letters into discourse units, we therefore decided to take the clause as a starting point. The procedure that was followed is discussed in this section.

As a first illustration, consider example (13) again, which is repeated below. In (13’), we present the same piece of discourse divided into discourse units.

(13) voorders mijn beminde soon soe laedt ick v l weten als dat mester ijacop
    further my beloved son so let I you know as that master IJacop
    blocken soon oock doot is ende soe heb ick oock soe daedlyck verstaen
    Blocken son also dead is and so have I also so now understood
    dat mester ijacop seluer doot is soo dat daer ael veel gestueruen zijn
    that Master IJacop self dead is so that there already many passed away are
    ‘Further, my beloved son, I let you know that Master IJacop’s son is also dead, and just now I have also understood that Master IJacop is dead himself, so that many have passed away’
(13')

a. voorders mijn beminde soon soe laedt ick v l weten
   ‘Further, my beloved son, I let you know’

b. als dat mester ijacop blocken soon oock doot is
   ‘that Master IJaap’s son is also dead,’

c. ende soe heb ick oock soe daedelyck verstaen
   ‘and just now I have also understood’

d. dat mester ijacop seluer doot is
   ‘that Master IJaap is dead himself,’

e. soe dat daer ael veel gestueruen sijn
   ‘so that many have passed away’

Each line in (13’) contains a clause, either an unambiguous main clause (a, c), an unambiguous subclause (b, d), or something in between (e). The clause in (13’e) begins with clause chaining soe dat (see Section 4), which may introduce new pieces of information, and therefore semantically main clauses, but the word order following the subordinator dat “that” is V-final, the word order typical of subclauses. What immediately stands out in (13’) is that any single discourse unit begins with a connective, which is a clear sign that connectives may fulfil functions similar to intonation contours in spoken language and to punctuation and capitalisation in written language, viz. chaining clauses by separating them.

Dividing the letters into discourse units, we dealt with various problems which will be illustrated here. We generally considered multiple subclauses dependent on one main clause as coordinated. If en was used as linkage, it was considered as verb phrasal coordinating en as in (18) above. But in case of a change of orientation or topic shift, en can often also be considered as a chaining connective as in (20) above. An example was given (13’e), where the subclause beginning with soe dat “so that” was informationally rather a new main clause. Now consider (22).

(22)

a. soo dient deese alleijn
   so serves this only

b. als dat gelieft
   as that pleases

c. aen alle beijde mijn neeuen corn en maghieltje de witte teseggen
   to both my cousins Corn and Maghieltje de Witte to-say

d. dat gistre haer ue mompijppe met suijcker
   that yesterday your barrel with sugar

e. dat selff gemaeckt hebbe
   which myself made have

f. gevolt hebbe,
   filled have
g. *en dat u e alsdan alsdaen lustigh candeel op mijn gesondt* and that you then all-together merrily caudle⁴ to my health *moet drinken* must drink

h. *en dat maghieltje candeel moet hebben sonder broot en corn* and that Maghieltje caudle must have without bread and Corn *met broot* with bread

‘This [letter] only serves to [ask you] to please tell both my cousins Corn and Maghieltje de Witte that yesterday I have filled your beer barrel with sugar, which I have made myself, and that you can all together merrily drink caudle to my health, and that Maghieltje must get caudle without bread, and Corn with bread’

From a hierarchical perspective, (22b) depends on (22a), (22c) depends on (22b), and (22d) depends on (22c). The clause in (22d) is intersected by a relative clause (22e), so that the clause is divided over two discourse units, viz. (22d) and (22f). The following (22g) may also depend on (22a), and the connective is therefore considered an instance of coordinating *en*. The piece of discourse in (22b–g), depending on the introductory formula (22a), concerns the fact that a barrel was filled with sugar, and that the addressees will therefore be able to drink caudle. In short: there is caudle. The next clause (22h) links up with this topic, especially with the act of drinking, but it also establishes a change of focus. It is no longer the existence of the caudle which is the point at hand, but the way in which it should be consumed by the people mentioned. The discourse unit has the form of another subclause, and it also begins with *en dat*, but because of the change of focus this connective is considered a case of chaining *en*.

Multiple main clauses linked with *en* are usually considered as chains, and *en* is a chaining connective in such cases. Naturally, exceptions occur when the main clauses are semantically entirely dependent on each other. Consider (23).

(23) a. *want sij ging 5 a 6 daeg daer na seijel* for they went 5 to 6 days there after sail+noun

b. *en sette al de hollanders aan lant sonder eete ofte eijts* and set all the Hollanders to land without food or anything

‘for they sailed away five or six days later, and set all the Hollanders ashore without food or anything’

The discourse unit in (23b) begins with *en*, which could have been omitted. The unit would have lacked a subject, but it should be noted that elision of pronominal subjects in such contexts is common in Early Modern Dutch letters and diaries. Still, these are not main clauses chained by bleached *en*. From the context it is clear that the

---

⁴. A warm drink made of wine, egg-yolk, sugar and cinnamon served with or without bread.
information in (23a) is only relevant in relation to the statement in (23b). There is a hierarchical structure between (23a) and (23b) in terms of information, with (23a) only providing background information to the more important statement in (23b). Therefore, *en* in (23b) is considered a case of coordination, not of chaining.

As became clear from (22e), we considered relative clauses as discourse units. These may intersect another clause that, in that case, becomes divided over different discourse units. Consider (24).

(24) a. *maer de liefde*
   but *the love*

   b. *die ghij tot mijn hebt*
   which you to my have

   c. *was doen ook wat grootter*
   was then also somewhat bigger

   ‘but the love you feel for me was also somewhat bigger then’

One could argue that (24a) and (24c) make up one discourse unit which would probably be the preferred analysis in hierarchical syntactic models. We take, however, a linear approach to syntax (Auer 2009; Sinclair & Mauranen 2006), assuming that it unfolds in actual discourse. Doing justice to this approach implies that (24c) is considered a separate discourse unit, which nevertheless clearly links up, also syntactically, with (24a). One could also argue that (24a) and (24b) constitute one unit, viz. the subject of the finite verb *was* “was” in (24c). Such an analysis, however, would obscure the fact that (24b) elaborates on a previously introduced notion, viz. the love.

As (24a) shows, in some cases discourse units are not clauses. Another example is the case of postponed (parts of) constituents, in particular when these appear after the final verb, that is after the sentence brace. See (25).

(25) a. *en wetet*
   and know+IMP

   b. *als dat ul oom heinderick Stoffelsz fijnstof den 7 januari 1672*
   as that your uncle Heinderick Stoffelsz Fijnstof the 7 January 1672
   *js overleden*
   is passed away

   c. *als mede hillegontie jans de dochter van jan Clasz*
   as also Hillegontie Jans the daughter of Jan Clasz
   ‘and be informed that your uncle Heinderick Stoffelsz Fijnstof passed away
   on 7 January 1672, as well as Hillegontie Jans, the daughter of Jan Clasz’

The information in (25c) is an extension of the subject of the subclause in (25b), viz. *ul oom heinderick Stoffelsz fijnstof*. It is a retraction (Auer 2009) in that it goes back to an existing constituent in order to modify it and therefore we consider (25c) not as a clause, but a discourse unit, due to its place outside the clause it retracts.
Likewise participial phrases, both with present and past participles, were regarded as separate discourse units. Consider (26).

(26) a. `zoo men bewust is'
   if man conscious is

b. `in wat situatie dat dezelve zig bevind,`
   in what situation that the same refl finds

c. `als eene enkelde Caper genoeg in staet zijnde`
   if a single privateer enough in position being

d. `om de Rivier in zijlende`
   to the river in sailing

e. `aen de wal komende`
   on the shore coming

f. `de Engelsche vlag te planten`
   the English flag to plant
   ‘if one would realise in what situation [the colony] finds itself, if a single privateer, sailing up the river, coming ashore, sufficiently being in a position to plant the English flag’

The discourse units in (26c, d, e) contain present participles. In accordance with traditional accounts, these are considered as adverbial clauses, and therefore as discourse units. Note, in passing, that (26f) is an infinite subclause, which we also considered as a separate discourse unit, since it contains its own predicate. Example (27) contains a past participle phrase.

(27) a. `wij zijn den darden dach ghe koome voor een plaets`
   we are the third day come for a place

b. `ghe naemt faro ofte kaep marij`
   called Faro or Cape Mary
   ‘the third day, we arrived at a place called Faro or Cape Mary’

Focus constructions, quite frequent in the letters, were also considered as separate discourse units. Usually they contain either a finite predicate (28a) or a present participle (29a).

(28) a. `wat dat den tyt hier aan gaat`
   what that the time here concerns

b. `dat is maar slecht`
   that is just bad

c. `daar word sterk van oorloog gesprook`
   there is strongly of war spoken
   ‘as to the times here, these are just bad. War is very much spoken about’
Apart from the verbal material in focus constructions, resumptives such as *dat* “that” in (28b) and *daar* “there” in (29b) also justify the focus constructions’ status as independent discourse units.

Vocatives such as *mijn beminde soon* “my beloved son” in (13’a) were not considered as separate discourse units, although this phrase interrupts the flow of information conveyed in the clause in which it is inserted. We nevertheless attached such vocatives to the discourse unit to which they appear closest or in which they are inserted as in (13’a).5

Formulaic opening passages were not taken into account. Consider example (30), where we divided the formulaic beginning of a letter into discourse units in accordance with the foregoing.

(30) a. *Eersame seer beminde lieve moeder ende broeders*  
honourable very beloved dear mother and brothers  

b. *Ick laet ul weten*  
I let you know  

c. *als dat Ick noch gesont ben*  
as that I still healthy am  

dl. *Ick hoop*  
I hope  

e. *dat het met ul en mijn twee broeders mede soo is*  
that it with you and my two brothers also so is  

f. *waer het anders*  
were it different

5. The reason is that we excluded formulaic parts, comprising salutations in the opening of letters. In many cases, these salutations take a similar form to vocatives. We therefore decided to incorporate vocatives into other discourse units, thus avoiding an imbalance of salutations and vocatives, with the first being omitted altogether, and the second constituting separate discourse units. Leaving out vocatives as well would have been another solution, but we decided to retain as much linguistic material as possible. Vocatives are often part of mainly non-formulaic passages and we excluded only full formulaic passages.
Split up as in (30), fixed formulae are spread over several discourse units. In Chapter 3, we argued that epistolary formulae were prefabs, that is, lexicalised units stored in and retrieved from memory as a whole. Therefore, we would rather consider (30d) and (30e) as one unit. Similarly, (30f), (30g) and (30h) constitute a unit. Thus, formulaic sequences appear to be a different kind of discourse, which is corroborated by the almost absence of chaining connectives.\(^6\) As stated above, we also singled out formulaic *en* as in (21). All clauses or constituents that are not part of the body of the letter, such as addresses, salutations, places and dates, and signatures were not taken into account. Salutations are often very well separated from the body of the letter, as in Figure 6, where the salutation *Monsr & freer Scheffens* “Mr & brother Scheffens” is placed above the body of the letter, almost as high as the place and the date on the upper right side.

There are, however, also many cases where the salutation is not separated from the body of the text at all, as in Figure 7. The first line reads *Beminde huijsvrouwe Fransyn-tje de Sorge* “beloved house wife Fransyntje de Sorge”, which is visually not distinguished from the following lines, whereas the writer used a different handwriting for the upper line with the place and the date. In cases such as Figure 7, the distinction

\(^6\) Moreover, there were strong social class and gender effects on the distribution of formulae (see Chapter 4). For instance, formulaic language occurs more often in the lower and lower-middle classes and in letters by women than in the upper and upper-middle classes and in letters by men. This implies that including formulaic clauses into the present analysis would create an imbalance: on the one hand, the number of units would rise in the case of language users employing epistolary formulae, while on the other hand, the proportion of units beginning with chaining *ende/en* would decrease, because this feature usually does not appear in formulaic passages.
between salutations and vocatives is blurred. One could argue that marked salutations are nothing other than vocatives that are visually separated from the body of the text. Since we excluded salutations from the analysis, we did not wish to put too much emphasis on vocatives either, and we therefore incorporated them into the closest discourse unit (see (13’)).

Contrary to formulaic opening passages, formulaic closing passages were incorporated into the data for the present analysis. The main reason is that the distinction between the middle part and the closing part is commonly far less sharply made than the distinction between a formulaic beginning and the middle part. The transition from the beginning to the middle part is usually marked by a formula such as voorts “further” (see Chapters 3 and 4), separating the formulaic opening passage from the non-formulaic middle part. Non-formulaic statements are virtually never inserted into formulaic beginnings. Towards the end of letters, formulaic and non-formulaic strings of words are, however, often intertwined. Furthermore, closing formulae may be followed by large pieces of discourse, and may also appear more than once. Consider (31), containing the discourse units just before the writer’s signature.

(31) a. voorts weet Jck op dit pas niet meer teschrijve
   further know I on this step not more to-write
b. als dat d heer valckenborgh u lieve man noch gesont was
   as that the Mr Valckenborgh your dear husband still healthy was
c. doe deese luyde vertrocke
   when these people left
d. vorder de groettenise aan sr heermans en zij huijs vrou
   further the greeting to Mr Heermans and his wife
e. soo hij noch te amsterdam is
   if he still in Amsterdam is
f. want men seijde
   for man said
g. dat hij hier koome soude
   that he here come would
h. *maer Jck conde dat niet geloove*  
   but I could not believe

i. *en ghij met u kindere zyt van mij van harte geroet*  
   and you with your children be\textit{IMP} from me from heart greeted

j. *en Jck vrsoeck*  
   and I request

k. *vrintschap met mijn huijs vrou te wille houde*  
   friendship with my wife to want keep

l. *en groet moer Jooris en allen goede vrinde en*  
   and greet\textit{IMP} mother Jooris and all good friends and  
   *bekende en onder andere mijn heer pergens*  
   acquaintances and amongst others Mr Pergens

‘furthermore, I do not know anything more to write on this occasion, except that Mr Valckenborgh, your dear husband, was still healthy when these people left. Further, greetings to Mr Heermans and his wife, if he is still in Amsterdam, for it is said that he would come here, but I could not believe that. And you with your children, be cordially greeted from me, and I request that you keep friendship with my wife. And greetings to Mother Jooris and all good friends and acquaintances and, amongst others, to Mr Pergens’

(31a) is a closing formula, which often rounds of the middle part and marks the transition to the closing part of the letter. Here, it rather introduces a non-formulaic piece of information which the writer apparently wants to incorporate in the letter (31b, c). In (31d), the writer again indicates the imminent closing of the letter by using a greeting formula. This, however, is followed by several discourse units conveying non-formulaic information on the person to be greeted (31e–h). Then follow three units, each introduced by chaining \textit{en}. The first (31i) could have been the final greeting, but something came to mind (31j, k), and again something came to mind (31l). Units (31i) and (31l) are formulaic strings, appearing many times in the letters (Chapter 3), while (31j,k) is a non-formulaic piece of information. What we learn from examples such as (31) is that indeed closing parts are less well separated from the middle part than beginnings, and that they often combine formulaic and non-formulaic sequences. Probably, when approaching the end of the letter, or the end of the piece of paper, writers once more realize what still needs to be said and add new information in units such as (31e–l). This discourse is in need of explicit organisation and this is what we see in (31d), and especially in (31i, j, l): chaining connectives are used for the organisation of newly thought of units.

We divided the subcorpora into discourse units along the lines explained above. This resulted in the following numbers for the subcorpora used in the present chapter. Table 4 gives the number of letters (N Le) and the number of discourse units (N DU) in the seventeenth-century subcorpus (across social class and gender). Table 5 presents
Table 4. Discourse units in the seventeenth century.

<table>
<thead>
<tr>
<th></th>
<th>LC</th>
<th>LMC</th>
<th>UMC</th>
<th>UC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Le</td>
<td>N DU</td>
<td>N Le</td>
<td>N DU</td>
<td>N Le</td>
</tr>
<tr>
<td>M</td>
<td>5</td>
<td>293</td>
<td>21</td>
<td>1,224</td>
<td>24</td>
</tr>
<tr>
<td>F</td>
<td>5</td>
<td>271</td>
<td>19</td>
<td>1,430</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>564</td>
<td>40</td>
<td>2,654</td>
<td>48</td>
</tr>
</tbody>
</table>

Table 5. Discourse units in the eighteenth century.

<table>
<thead>
<tr>
<th></th>
<th>LC</th>
<th>LMC</th>
<th>UMC</th>
<th>UC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N Le</td>
<td>N DU</td>
<td>N Le</td>
<td>N DU</td>
<td>N Le</td>
</tr>
<tr>
<td>M</td>
<td>24</td>
<td>782</td>
<td>24</td>
<td>849</td>
<td>24</td>
</tr>
<tr>
<td>F</td>
<td>2</td>
<td>95</td>
<td>14</td>
<td>723</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>26</td>
<td>877</td>
<td>38</td>
<td>1,572</td>
<td>48</td>
</tr>
</tbody>
</table>

the number of letters (N Le) and the number of discourse units (N DU) in the eighteenth-century subcorpus (across social class and gender). In both periods, the letters make up well over 8,000 discourse units.

6.2 Sociolinguistic patterns

The case study presented here focuses on the distribution of chaining *en* across social class and gender in the seventeenth- and eighteenth-century subcorpora. As stated above, we expect this more oral feature to be relatively more frequent in letters by writers who are less involved in the written culture. The hypothesis also follows from Culpeper & Kytö (2010: 165–167, 178–179), who found the highest proportion of what they call clausal *and* in Early Modern English speech-based genres such as trials, drama and witness depositions, as well as in history books. The latter result they explain by the narrative nature of history, where the events that make up the narrative are sequenced by clausal *and*. They also suggest a diachronic decrease of the use of chaining *en*.

In order to establish possible variational patterns, we counted the number of discourse units (N DU) beginning with chaining *en* (N *en*) per social class and per gender. The results for the seventeenth-century subcorpus are given in Table 6. As Table 6 shows, 1,109 discourse units out of a total number of 8,282 discourse units begin with chaining *en*, which comes down to 13%. Put differently, about 1 in every 7.5 discourse units begins with chaining *en*. In Figure 8, we plotted the proportion of discourse units beginning with clause chaining *en* per social class and per gender. As can be seen from
Chapter 7. Clause chaining between spoken and written language

Table 6. Discourse units with *en* in the seventeenth century.

<table>
<thead>
<tr>
<th>17th c.</th>
<th>LC</th>
<th>LMC</th>
<th>UMC</th>
<th>UC</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N DU N en</td>
<td>N DU N en</td>
<td>N DU N en</td>
<td>N DU N en</td>
<td>N DU N en</td>
</tr>
<tr>
<td>M</td>
<td>293  48</td>
<td>1,224 202</td>
<td>1,637 194</td>
<td>1,230 104</td>
<td>4,384  548</td>
</tr>
<tr>
<td>F</td>
<td>271  54</td>
<td>1,430 198</td>
<td>1,528 210</td>
<td>669   99</td>
<td>3,898  561</td>
</tr>
<tr>
<td>Total</td>
<td>564 102</td>
<td>2,654 400</td>
<td>3,165 404</td>
<td>1,899 203</td>
<td>8,282 1,109</td>
</tr>
</tbody>
</table>

Figure 8. Proportion of discourse units beginning with *en* in the seventeenth century, by social class and gender.

In Table 7, we present the results for the eighteenth-century subcorpus.
Table 7. Discourse units with *en* in the eighteenth century.

<table>
<thead>
<tr>
<th></th>
<th>LC</th>
<th></th>
<th>LC</th>
<th></th>
<th>UC</th>
<th></th>
<th>UC</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>DU</td>
<td>N</td>
<td>DU</td>
<td>N</td>
<td>DU</td>
<td>N</td>
<td>DU</td>
<td>N</td>
<td>DU</td>
</tr>
<tr>
<td>M</td>
<td>782</td>
<td>123</td>
<td>849</td>
<td>133</td>
<td>1,125</td>
<td>107</td>
<td>1,609</td>
<td>112</td>
<td>4,365</td>
<td>475</td>
</tr>
<tr>
<td>F</td>
<td>95</td>
<td>4</td>
<td>723</td>
<td>83</td>
<td>1,788</td>
<td>217</td>
<td>2,016</td>
<td>164</td>
<td>4,622</td>
<td>468</td>
</tr>
<tr>
<td>Total</td>
<td>877</td>
<td>127</td>
<td>1,572</td>
<td>216</td>
<td>2,913</td>
<td>324</td>
<td>3,625</td>
<td>276</td>
<td>8,987</td>
<td>943</td>
</tr>
</tbody>
</table>

As Table 7 shows, 943 discourse units out of a total number of 8,987 discourse units begin with chaining *en*, which amounts to 10.5%. On average, 1 in every 9.5 discourse units begins with clause chaining *en*. In Figure 9, we plotted the proportion of discourse units beginning with clause chaining *en* across social class and gender. Again, the black total columns show a steady decrease in the use of clause chaining *en*, from 14.5% in the LC to about 7.5% in UC, that is, from 1 token in every 7 discourse units to 1 in every 13 discourse units. A similar decrease can be witnessed in the dark grey columns for male letter writers. As in the seventeenth century, the male scores are quite similar in the LC and the LMC, viz. almost 16%, and then sharply drop in the UMC and the UC. The light grey columns representing letters by female writers also

---

7. The difference with the seventeenth-century results, where 1109 out of 8282 discourse units begin with chaining *en* (Table 6), is statistically significant according to a two-tailed Chi-square test with Yates correction (p<0.0001).
drop, in particular from the similar scores in the LMC and the UMC, viz. about 12%, to 8% in the UC. Note that from 24 letters by UC men, no fewer than 6, so 25%, have no tokens of clause chaining *en* at all. In all other social/gender categories, in the seventeenth as well as in the eighteenth century, letters without a single token only occasionally occur. The one score seemingly out of place is the extremely low proportion of 4% for LC women. It should be remembered that LC women are represented by only two letters, which means that the score for LC women must be treated with appropriate caution. As in the seventeenth century, the overall picture shows a decrease of chaining *en* as the social class index increases, with possible gender effects in that women use more chaining *en* than men. In any case, we would expect such a gender effect, considering the results for the seventeenth century as well as for formulaic language (Chapter 4), but with a problematic score for women in the LC, and with LMC women scoring less than LMC men, there is no conclusive evidence for a gender distribution. As in the seventeenth century, however, the highest scores are in the LC and the lowest score is for UC men, who on average use 1 token of chaining *en* in every 14 discourse units.

Turning to the diachronic results now, we plotted the total columns from Figures 8 and 9 together in Figure 10. In both the seventeenth (the black columns) and the eighteenth century (the light grey columns), there is a steady decrease from LC to LMC, UMC and UC. Within the social classes, seventeenth-century total scores are consistently higher than eighteenth-century total scores. Therefore, we conclude that the use of clause chaining *en* decreases over time and that letter writers from the UC, and UC men in particular (see Figures 8 and 9), lead this change.

![Figure 10. Proportion of discourse units beginning with *en* in the seventeenth and eighteenth centuries, by social class.](image-url)
In sum, the results we found are in line with the hypotheses formulated at the beginning of Section 6 and 6.2. Clause chaining *en* is a fairly oral feature which frequency decreases as the informants are more involved in the written culture. Clause chaining *en* is more prevalent in lower classes than in upper classes, and, with the necessary reservations, also more among female writers than among male writers. Moreover, the use of clause chaining *en* decreases from the seventeenth to the eighteenth century, paralleling the increasing textualisation of society in general.

7. The interplay of punctuation and clause chaining

In Section 5, we observed that the rise of punctuation in private letters was a gendered change from above. In Section 6, it has become clear that the decrease in the use of clause chaining *en* was a possibly gendered change from above as well. In this section, we will discuss the interplay of these changes.

We should point out beforehand that the presence of punctuation does not self-evidently imply the absence of bleached connectives. First, the use of semantically bleached clause chaining connectives is still widespread in spoken language (cf. the work by Chafe), and very probably also in conceptually oral genres such as private email or chat room language. Second, writers in our letter corpora may combine strategies, that is, use bleached connectives in one place, and punctuation in another, as in (32).

(32) a. *als dat uw Eedl bij uw vader is,*
   as that your Honour at your father is

   b. *het welke mij lief is,*
   the which me+DAT dear is

   c. *al uw maar wel Accordeeren Kont met uw suster Lott*
   if you only well agree can with your sister Lott

   d. *en ik hebbe gehoorht*
   and I have heard

   e. *als dat de oude man nog gesont is*
   as that the old man still healthy is
   ‘that your Honour is at your father’s, which is dear to me, if only you could get along with your sister Lott. And I have heard that the old man is still healthy’

Whereas discourse units (32a) and (32b) are separated from the following unit by a comma, (32d) is linked to (32c) by clause chaining *en*. A third reason why punctuation and bleached connectives do not need to be communicating vessels is that one of the four strategies described in Section 2 involved bleached connectives as well as punctuation, not as in (32) in a larger stretch of discourse, but as in (33) in one and the same position between two discourse units.
(33) a. *doe was hij opperstuurman op 't ygenste schip Amsterdam*, then was he first mate on the same ship Amsterdam

b. *en wij bennen den 3 Maaij in Ree gekomen op Caap* and we are the 3 May in roadstead come+part on Cape ‘then he was first mate on the same ship Amsterdam, and 3 May we arrived in the roadstead at the Cape’

For these reasons, the rise of punctuation and the decline of clause chaining *en* need not be intertwined, and if they are, it will be difficult to observe that when focusing solely on the results per social category. In this section, we therefore take a closer, more qualitative look at a number of letters.

We can be brief about seventeenth-century female writers, since almost none of them use punctuation. In the eighteenth-century subcorpus used here, we also find no straightforward correlation between punctuation and clause chaining *en*, since letters with punctuation also contain clause chaining *en*. The same applies to letters by LC and LMC men in both the seventeenth and the eighteenth century: letters with punctuation commonly also have bleached connectives. Clearly, punctuation and clause chaining *en* are not alternative strategies here. Rather, it seems that punctuation is brought in as an additional means of clause linking in these letters by women, and by men from the LC and the LMC. The letter writers maintain the relatively oral use of chaining *en*, incorporating the more literate means of clause chaining through punctuation at the same time.

The case of the 24 seventeenth-century letters by UMC men, however, differs. As, on the one hand, 14 (out of 24) have punctuation and all of these also exhibit bleached clause chaining *en*; on the other hand, the three letters with the highest proportion of clause chaining *en*, viz. 49%, 26% and 23% respectively, lack punctuation altogether. This suggests that on an oral–literate continuum with regard to the use of chaining *en*, the letters closest to the oral pole are also farthest removed from the literate pole with regard to punctuation. This becomes even more plausible when focusing on UC men in the seventeenth century. Only three out of 15 letters are without punctuation. At the same time, two of these have the highest scores of clause chaining *en*, viz. 23% and 15%, respectively. Moreover, the letters in question are rare examples of letters by UC men with fully formulaic beginnings, as exemplified in example (1) in Chapter 3. These two writers are both relatively young men, and members of the upper class by birth: Willem Waterdrinker, a captain’s son, was a 24-year-old shipmaster and Doede Ennes Star, a Rear Admiral’s son, was a hireling of about 20 years. While we cannot exclude the possibility that they needed to be able to read and write in their profession, there does seem to be a difference from the social profile of the three seventeenth-century UC men with the lowest proportion of clause chaining *en*. These three writers, who use punctuation and hardly any bleached connectives, are Hieronijmus Sweerts, a poet and printer/publisher (with only 2% of his discourse units beginning with chaining *en*), Henricus Cordes, a minister in training (4%), and Anthonius Scherius, a minister
Summing up, we have two writers with an occupation fairly different from other UC men in that it is not an unambiguously literate profession, and these two writers produce formulaic passages atypical of UC men, as well as high numbers of clause chaining *en*, while their letters lack punctuation.

Turning to UMC and UC men in the eighteenth century, less obvious results are obtained. Only seven out of 24 letters by UMC men have no punctuation and these are not the seven letters with the highest proportion of clause chaining *en*, though six are above the average UMC men score of 9.5% of discourse units beginning with *en*. So the correlation is less outstanding than in the seventeenth century. Furthermore, five out of 24 letters by UC men lack punctuation, while a possible link with the presence of bleached *en* cannot be made.

In letters from the lower classes, and in letters by women, no straightforward evidence is found for an interplay of punctuation and clause chaining *en*. The bulk of these letters contain clause chaining *en*, while some of them introduce punctuation as just another means of clause chaining. Thus, punctuation does not exclusively rise at the expense of clause chaining *en*. The results suggest that the use of punctuation first arose as an additional strategy for clause linking. In the upper classes, however, and in letters by men in particular, there is some evidence that the presence of punctuation and the absence of clause chaining *en* are correlated. High proportions of clause chaining *en* are in some cases combined with lack of punctuation. Moreover, formulaic language prevails in such cases. Building on the idea that the use of formulae was linked to writing experience (Chapters 4 and 5), and taking into account the writers’ occupations, we suggest that writing experience as a function of everyday life probably influenced the use of punctuation and clause chaining *en*.

8. Conclusions

In this chapter, we discussed the way in which letter writers chained clauses in order to create larger pieces of discourse. Different strategies were in use, some more oral, others prototypically written. An investigation of these strategies enables us to locate the genre of seventeenth- and eighteenth-century private letters on an oral–literate continuum. The analyses in the present chapter therefore link up with the approach taken in Chapter 2 on the degree of orality in the field of orthography, phonology and morphology.

On the more oral side of the oral–literate continuum, we found semantically bleached connectives grammaticalising into discourse markers. Clause chaining *en* was the most frequent one of these. Thus, the relative orality of written genres can be measured by the extent to which bleached connectives are employed for clause chaining. Culpeper & Kytö (2010: 165, 167, 178–179) have shown that *and* occurs more often in Early Modern English speech-based genres such as trials, drama and witness depositions than in scientific texts, and have also suggested that its frequency decreases over time. A diachronic decrease would indicate a move along the oral–literate
continuum from the oral pole towards the literate pole.\textsuperscript{8} This ties in with what we observed in the present chapter. Different variables affected the distribution of chaining *en*, in that this more oral feature was more in use in the lower classes than in the upper classes, more among women than among men, and more in the seventeenth than in the eighteenth century. At the same time, punctuation as the prototypically written means of clause chaining was introduced into the letters by men from the UMC and the UC. The rise of punctuation and the decline of clause chaining *en* were probably gendered changes from above, which rendered the language of letters less oral and more literate. At the level of clause chaining, therefore, we may say that the genre of private letters was less oral in the eighteenth century when compared to the seventeenth century, and that this change from oral to literate was largely led by upper and upper-middle class men.

\textsuperscript{8} Such a development would also be in line with Biber & Finegan (1989), who claim that three English genres (fiction, essays and letters) have gradually become more oral from the seventeenth to the nineteenth century, with however a temporary move away from orality in the eighteenth century.
Chapter 8

Variation and change in the relative clause

1. Introduction

In the preceding chapters, we discussed a variety of topics commonly associated with historical sociolinguistic research on private letters, such as the degree to which ‘oral’ elements can be found in the written code, and the use of epistolary formulae and terms of address. In Chapters 8, 9 and 10, we will focus more generally on morphosyntactic features undergoing change in seventeenth- and eighteenth-century Dutch. These features have been selected for their relatively high frequency in our letter corpora. Our sociolinguistically stratified letter corpora offer new ways of analysing these changes, and our main aim therefore is to establish possible sociolinguistic patterns in the spread of the changes at hand. Nevertheless, internal factors have also to be taken into account in each of these changes. The first morphosyntactic feature, to be discussed in the present chapter, concerns relative (pronominal) adverbs. Referring to the initial consonants, these change from *d*-forms into *w*-forms.

In Section 2, we will introduce the relativisation changes in more detail, and briefly review previous research. We will approach the topic from the perspective of syntactisation, which leads us to several hypotheses about the sociolinguistic distribution of *d*- and *w*-forms (Section 3). Section 4 discusses the diachronic, social and gender distribution found in the corpora. In Section 5, we link the concept of syntactisation to writing experience, and explain the sociolinguistic effects established in Section 4 from this perspective. Since quite a few epistolary formulae, which are so characteristic of our material, contain relative clauses, we discuss in Section 6 the interplay of formality and the relativisation changes. This discussion raises questions about the importance of clause integration, which will be the topic of Section 7.

2. Relativisation in the history of Dutch

In the seventeenth and eighteenth centuries, one of the significant changes in the grammar of Dutch concerns the form of relativisers. In general, relative pronouns, adverbs and pronominal adverbs change from a *d*-form into a *w*-form.¹ This change

¹ Note that our description relates to written, supraregional varieties of Dutch. In present-day dialects, much more variation is found, including *d*-forms still used in examples such as (6) and (7) below; see DynaSAND (2006).
from *d*- to *w*-relativisation constitutes a major shift in the grammar of Dutch, as in other Germanic languages (Rissanen 1999: 292–301; von Polenz 1994: 278–279). In the history of Dutch, the change began in the late Middle Dutch period, probably in the fourteenth or fifteenth century (van der Horst 2008: 603, 703), and is not yet completely finished. It affects any kind of relative clause (restricted and appositive relative clauses, including continuative relative clauses), any kind of relativiser (pronouns, adverbs and pronominal adverbs), and any kind of syntactic/semantic context (dependent and independent or free relative clauses). A few examples will illustrate this. In (1–4), the changes in the pronominal system are shown with free relatives (1, 2) and with nominal antecedents (3, 4). The (a)-examples are Middle Dutch, the (b)-sentences are Modern Dutch. In examples (1) and (3), the antecedent is inanimate, in (2) and (4) it is animate. The change represented by (1, 2) is complete. The change in (3) is in progress, with the *w*-form being common in colloquial Dutch, while the *d*-form is preferred in the written standard. Only few speakers would accept (4b), but *w*-forms do appear in this position, including in written language.

(1) a. *Had ic ghevonden dat ic zoek*  
Had I found REL I seek  
‘Had I found what I was looking for’

b. *Na 5 weken had ik gevonden wat ik zocht*2  
After 5 weeks had I found REL I sought  
‘After five weeks, I had found what I had been looking for’

(2) a. *Die sine cuusheit uerlieset, die uerlieset sine siele*  
rel his chastity loses that loses his soul  
‘He who loses his chastity, loses his soul’

b. *wie zijn KUISHEID bewaakt mag door elk deur die hij/zij*  
rel his chastity guards may through each door that he/she  
*wil het paradijs binnentreden!*3  
wants the paradise enter  
‘He who guards his chastity, may enter paradise through any door he/she wants to’

(3) a. *dat woordt dat die heilighe man job sprac*  
that word REL that holy man Job spoke  
‘the word that the holy man Job spoke’

b. *Neger, ja, dat is het woord wat Totti tegen mij zei*4  
negro yes that is the word REL Totti to me said  
‘Negro, yes, that is the word that Totti said to me’

2. http://www.datingwebsites.nl/reviews/second-love/
(4) a. **vrouwen, die ter merct brengen wouden eyer ende botter**
    (van der Horst 2008: 601)
    women rel to-the market bring would eggs and butter
    ‘women, who wanted to bring to the market eggs and butter’

b. **de grote minderheid van de vrouwen wie op die manier gebruik**
    the big minority of the women rel on that way use
    ‘the vast minority of women who make use of the internet in that way’

Similar changes have affected relative adverbs and relative pronominal adverbs (5–7).
In (5), the change in the free relative adverb is shown. In (6), the antecedent is an inanimate noun and the relativiser is the locative adverb. Example (7) contains pronominal adverbs, daer + in ‘therein’ and waarin ‘wherein’, respectively.

(5) a. **Sine es niet daer si was te voren**
    (van der Horst 2008: 477)
    she is not rel she was before
    ‘She is not where she was before’

b. **ze is niet waar ze eerder was**
    she is not rel she earlier was
    ‘She is not where she was earlier’

(6) a. **tot Brusel, daer sy hoer antwoord kreghen**
    (van der Horst 2008: 703)
    in Brussels rel they their answer got
    ‘in Brussels, where they got their answer’

b. **te Brussel, waar zij haar debuut maakte**
    in Brussels rel she her debut made
    ‘in Brussels, where she made her debut’

(7) a. **den viere / daer die bouc in bernende lach**
    (van der Horst 2008: 498)
    the fire rel the book in burning lay
    ‘the fire in which the book lay burning’

b. **het vuur waarin ze branden zal niet doven**
    the fire rel they burn shall not smother
    ‘the fire in which they burn will not smother’

In all these cases, *d*-relativisers are giving way or have given way to *w*-relativisers. We may also argue that interrogatives replace demonstratives as the main means of

5. http://www.wowforum.nl/viewtopic.php?f=16&t=18006&st=0&sk=t&sd=a,
8. http://www.allaboutworldview.org/dutch/bestaat-de-hel.htm
relativisation, as the $d$-relativisers are identical to – and probably originate from – demonstratives, while $w$-forms are identical to interrogatives. As stated before, the change is complete in the case of (pronominal) adverbs (cf. footnote 1) and the crucial period in this change was the seventeenth and eighteenth centuries (e.g. van der Wal 2002). The relative (pronominal) adverbs therefore represent an interesting case for the present study, which focuses on exactly these periods. In this chapter, we will consider the distribution of $d$- and $w$-forms in our letter corpora from the perspective of both language variation and change.

The change from $d$- to $w$-forms has gained a lot of attention, in the literature both on English⁹ and on Dutch. Studies on Dutch such as van der Horst & Storm (1991), van der Horst (1993), de Schutter & Kloots (2000), Schoonenboom (1997) and van der Wal (2002) have argued that a crucial factor in the spread of $w$-forms was the (in)definiteness of the antecedent. $W$-forms would first appear in indefinite contexts, and only later on in definite contexts. The contexts in question are, from highly indefinite to highly definite:¹⁰

I free relatives

a. *je moet doen wat ze zegt* ‘you must do what she says’
b. *je moet eten waar ze gedekt heeft* ‘you must eat where she has set the table’

II the antecedent is an entire sentence

a. *ze deed erg haar best, wat we heel flink vonden* ‘she really did her best, which we thought was very plucky’
b. *ze deed erg haar best, waar we heel blij om waren* ‘she really did her best, where we were very glad about’

III the antecedent is a word such as iets ‘something’, niets ‘nothing’, alles ‘everything’, veel ‘much’, weinig ‘little’, genoeg ‘enough’

a. *alles wat hij wil* ‘everything that he wants’
b. *is er iets wat/dat ik voor je kan doen?* ‘is there anything that I can do for you?’
c. *overal waar we komen* ‘everywhere where we come’
d. *alles waar ik naar verlang* ‘everything that I long for’
e. *is er iets waar je niet van houdt?* ‘is there something that you do not like?’

IV the antecedent is a substantivised adjective, usually a superlative

a. *het vriendelijkste wat/dat Jan gedaan heeft* ‘the friendliest [thing] that John has done’
b. *het gekste wat/dat ik ooit heb meegemaakt* ‘the craziest [thing] that I have ever experienced’

⁹. For references, see Rissanen (1999) and Bergs (2005).

¹⁰. The examples are in part taken from van der Horst (1988); *dat/wat* indicates that both forms are accepted in modern Dutch, *dat* (wat) indicates that the first form is prescribed in the present-day standard, while the second is also very common.
c. *het grootste waar ik ooit ben geweest* ‘the biggest [e.g. country] where I have ever been’

d. *het gekste waar je van houdt* ‘the craziest [thing] that you like’

V the antecedent is an undefined NP

a. *een boek dat (wat) ik mooi vind* ‘a book that I like’

b. *een boek waar ik graag in lees* ‘a book where I like to read in’

c. *een land waar ik graag kom* ‘a country where I like to come’

VI the antecedent is a definite NP

a. *het boek dat (wat) ik gisteren las* ‘the book that I read yesterday’

b. *het boek waar ik gisteren in las* ‘the book where I was reading in yesterday’

c. *het land waar ik was* ‘the country where I was’.

Justification for this hierarchy of definiteness or indefiniteness comes from the fact that the new w-forms are originally interrogatives, and therefore indefinite. Rutten (2010) argues that this approach, despite its intuitive probability, cannot explain all the empirical facts. There are, for example, seventeenth- and eighteenth-century diaries and letters with d-forms throughout, and with w-forms in context III, or in contexts I and III, or even only in contexts III and VI (Rutten 2010: 4). Rutten (2010) furthermore argues that w-forms first occur in specific constructions, and are then generalised to a more abstract level, where w becomes a formal marker of relative clauses. Since clauses can also be connected paratactically with demonstratives, the adoption of a marker specifically for relative clauses implies that the difference between paratactic and hypotactic structures becomes more explicitly coded in the linguistic system. Clauses with a w-form are necessarily subordinate clauses – at least syntactically, since subordinate relative clauses can be semantically coordinated in so-called continuative relative clauses. In the approach taken in Rutten (2010) and Rutten & van der Wal (to appear) the relevance of a strongly coded difference between parataxis and hypotaxis is presupposed. When hypotactic structures become more explicitly coded, the language becomes more grammatical, and more syntactic. As such, the language grammaticalises, or syntacticises. In the present study, we continue this line of research, providing evidence that changes in the relative clause can indeed be explained in terms of syntacticisation. We discuss this approach in more detail in the next section.

3. **Syntacticisation of the relative clause – hypotheses**

A frequently used basic working hypothesis in evolutionary linguistics is that syntax developed later in the evolution than simple signs and words (Bickerton 1990; Nowak & Krakauer 1999; Nowak, Plotkin & Jansen 2000; Jackendoff 1999; Tomasello 2008), recalling Givón’s well-known dictum, viz. that language develops from discourse to
grammar, a development which he termed syntacticisation (1979: 208). By this, Givón meant, first, that human pragmatic and semantic operations, including meaning-making through words, precede encoding into syntactic structures, and second, that basic syntactic structures may become more syntactic over time, even though syntactic structures may, in their turn, erode over time (1979: 209). Givón (2009: 10) presents a three-step evolutionary model:

1. single words > simple clause
2. simple clause > clause chains (parataxis)
3. clause chains > complex/embedded clauses (syntax).

Steps (2) and (3), which Givón labels as the transition from parataxis to syntax, have often been described as a development from parataxis through hypotaxis to subordination (Hopper & Traugott 2003: 177; cf. Heine & Kuteva 2007: 210–261). Here, parataxis refers to independent and unembedded clauses, hypotaxis to dependent but unembedded clauses, and subordination to dependent and embedded clauses (Hopper & Traugott 2003: 178). Such steps are at the core of our idea of syntacticisation, and we will interpret changes in the Dutch relative clause from this perspective. The change from *d-* to *w*-forms presents a case of syntacticisation, and as such it is part of a larger trend in the linguistic system towards more hypotactic structures.11

Importantly, whereas the development from parataxis to hypotaxis/subordination may be a general trend in linguistic systems (cf. Givón), it could equally well be socially and/or culturally motivated, especially from an evolutionary perspective (Croft 2000). If language is considered as an evolutionary system, which adapts to social/cultural circumstances, it can be said to have been enormously impacted by the development of literacy. Already in 1983, Pawley & Syder formulated their ‘adaptation hypothesis’ (cf. also Becker et al. 2009 and the references there):

> Our principal hypothesis is that in the history of English certain usages have developed or gained preference in a given system because they are advantageous in the circumstances. We are dealing with an ecology of grammar, in which forms of construction are molded to suit the constitutive conditions and purposes of face-to-face talk, on the one hand, and impersonal written communication on the other (Pawley & Syder 1983: 552).

The basic idea is that the social/cultural context in which the language is used influences its grammar. One of the most preponderant aspects of this context is the channel: is the language spoken or written? Pawley & Syder (1983: 557–558) list systematic

11. If this brief overview gives the impression of increasing complexity as a general trait of human language evolution, it should be noted that simplification also occurs, which, however, does not usually mean the reversal of evolved syntactic structures, but prototypically involves verbal and nominal deflection or loss of inflection (Trudgill 2011; cf. also Dahl 2004; Sampson, Gil & Trudgill 2009). Deflection often co-occurs with syntacticisation: for instance, when Dutch and English lost most of their cases, more prepositional phrases developed and word order became more rigid (e.g. Lass 1999: 138–140).
differences between written and spoken communication, which are, for instance, also well-known from the work of Chafe (1985, 1994), and which are at the heart of corpus-based research into genre differences (Biber & Conrad 2009; cf. also Koch & Oesterreicher 1985). Whereas discourse phenomena may be coded in gestures, pauses, intonation and facial expressions, the written channel needs other means to code pragmatic meanings, one of which may be employing syntax, as we will argue in this chapter. Pawley & Syder (1983: 558), for that matter, note that personal letters may cross the divide between prototypically conversational discourse and prototypically autonomous discourse, that is, between speech and writing.

In terms of syntactic complexity, the written channel promotes hypotaxis more than the oral channel does. The more people write, the more the language becomes written-like, that is, more syntactic. It has been noted that written German from the sixteenth century onward is characterised by an explosion of coordinators and especially of subordinators (von Polenz 2000; Nitta 2002). The number of clause-linking devices rose, as the language was more often used for expressing meanings that had previously been coded in Latin. Also, many more people had access to literacy. The increase in subordinators has been established for Late Middle and Early Modern Dutch as well (van der Horst 2008: 983–992). In the domain of clause linkage, the languages were becoming more syntactic, and we consider the changes in the relative clause as part of this development.

It should be stressed that Dutch relative clauses were not paratactically linked to antecedents before the change from d- to w-relativisation. Consider example (6a) again, and its construed paratactic counterpart, here presented as (8a) and (8b):

(8) a. tot Bruesel, daer sy hoer antwoort kreghen
    in Brussels, rel they their answer got
    ‘in Brussels, where they got their answer’

b. tot Bruesel, daer kreghen sy hoer antwoort
    in Brussels, dem got they their answer
    ‘in Brussels, there they got their answer’

The subordinate relative clause attached to the noun Bruesel in (8a) has the finite verb in final position, signalling that the clause is subordinate. In (8b), a main clause with V2 follows the prepositional phrase tot Bruesel, characterised by inversion of the subject sy and the finite verb kreghen. In short, there was a word order difference between subordinate clauses and main clauses, and therefore between relative clauses and main clauses. The development of w-relativisers, and the replacement of there by where in...

---

12. Note that Pawley & Syder (1983: 566–567) do not adhere to the idea that speech is more coordinated and written language more subordinated. They describe the difference rather in terms of the level of integration of subordinate clauses into main clauses. For critical assessment of the idea that written language is more hypotactic than spoken language, see e.g. Beaman (1984) and Auer (2002).
examples such as (8a), therefore does not mean that relative clauses are suddenly syntactically marked, but merely that they become more strongly marked in the linguistic system when they obtain the $w$ as an additional distinguishing characteristic. The change from $d$- to $w$-relativisers thus does not map the change from parataxis to hypotaxis, but from hypotaxis to more hypotaxis.

It should be noted, however, that the difference between V2 main clauses and V-final subclauses was not as firm in the Middle Dutch and Early Modern Dutch periods as it is in present-day standard Dutch. We should keep in mind that the development of mandatory V2 and V-final subclauses takes place in the period from the fourteenth to the seventeenth centuries, so from Late Middle Dutch to Early Modern Dutch, when textuality spreads, and when the language is moulded so as to perform more written functions. Therefore, Burridge (1993) interprets the development of mandatory V2 in main clauses as well as mandatory V-final in subclauses as a result of grammaticalisation, by which she means what we call syntacticisation. Our approach to the relative clause thus ties in with the work of Burridge (1993).

Against the background of the discussion so far, the following hypotheses can be formulated. If the syntacticising changes in the relative clause are dependent on social or cultural circumstances such as increased literacy and the ongoing textualisation of society in general, we expect the change from $d$- to $w$-relativisation to be related to writing experience. If the written channel promotes hypotaxis, people who are more involved in the written culture and who use the written code more frequently are expected to use the prototypically written forms more frequently. In other words, we expect the change towards $w$-relativisation to be a change from above. In terms of writing experience, we expect clear social and gender distributions, apart from the diachronic development towards $w$-forms.

4. Variation and change in the relative clause – case studies

In order to test the hypotheses formulated in the previous section, we carried out a series of case studies on the seventeenth-century autograph subcorpus and on the eighteenth-century subcorpus (cf. Chapter 4, Section 3.1). In the present section, we report on these case studies and present the diachronic, social and gender patterns found (4.1, 4.2). Finally, we conclude that the change from $d$- to $w$- was indeed a change from above (4.3).

4.1 Extracting $d$'s and $w$'s – diachronic results

All relative adverbs and relative pronominal adverbs with initial $d$ and $w$ along with their syntactic context were extracted from the two subcorpora. Some ambiguous examples were excluded, of which, due to the absence of a finite verb, it was impossible
to establish their status as either main clause or subclause. In some cases, a *d*-form such as *daer* ‘there’ could also be interpreted as a conjunction fulfilling an argumentative function, usually as the concessive subordinator *daer* ‘although, whereas’, which is a grammaticalised result of the relative adverb. Consider example (9) with the free relative *daer* that contains an ‘implied antecedent’ and could therefore be interpreted as *there where*, but also as a concessive.

(9) god ... dancken ende loeuen voor de genaede die heij aen ons beweijst
god ... thank and praise for the mercy that he to us shows
*daer*  weij sulcke kinderen van verderf sijn
rel / conj we such children of doom are
‘[we should] thank God and praise Him for the mercy which He shows to us there where/although/even though we are such children of doom’

The concessive interpretation may be favoured, but the original adverbial interpretation of the locative relative adverb *daer* remains possible: it is in the metaphorical position of doom in which man finds himself that God nevertheless shows mercy to him. In cases such as (9), *daer* allows the grammaticalised argumentative function but can also be interpreted as the relative adverb and therefore such examples were not excluded from the results.

For the seventeenth century, this procedure resulted in 269 relative clauses, 150 of which beginning with a *d*-relativiser, and 119 with a *w*-relativiser, which amounts to 56% *d*-forms and 44% *w*-forms respectively. In the eighteenth-century subcorpus, we found 598 relative clauses, 169 of which with a *d*-forms, and 429 with a *w*-form, amounting to 28% *d*-relativisation and 72% *w*-relativisation. See Table 1.

### Table 1. Relativisation in the seventeenth and eighteenth centuries.

<table>
<thead>
<tr>
<th></th>
<th><em>d</em>-relativiser</th>
<th></th>
<th><em>w</em>-relativiser</th>
<th></th>
<th>Total</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>17th century</td>
<td>150</td>
<td>56</td>
<td>119</td>
<td>44</td>
<td>269</td>
<td>100</td>
</tr>
<tr>
<td>18th century</td>
<td>169</td>
<td>28</td>
<td>429</td>
<td>72</td>
<td>598</td>
<td>100</td>
</tr>
</tbody>
</table>

Clearly, the corpus results confirm the ongoing spread of the incoming *w*-relativisers, rising from less than fifty per cent in the 1660s/1670s to over seventy per cent in the 1770s/1780s.

### 4.2 Social and gender variation

After having established the diachronic difference, we studied possible patterns linked to the social class index. Table 2 presents the distribution of *d*- and *w*-relativisers in absolute numbers across social class in the seventeenth and eighteenth centuries.
Table 2. Relativisation in the seventeenth and eighteenth centuries, by social class.

<table>
<thead>
<tr>
<th>Social Class</th>
<th>17th century</th>
<th>18th century</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>d</em>-relativiser</td>
<td><em>w</em>-relativiser</td>
</tr>
<tr>
<td>LC</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>LMC</td>
<td>29</td>
<td>20</td>
</tr>
<tr>
<td>UMC</td>
<td>89</td>
<td>74</td>
</tr>
<tr>
<td>UC</td>
<td>22</td>
<td>23</td>
</tr>
<tr>
<td>Total</td>
<td>150</td>
<td>119</td>
</tr>
</tbody>
</table>

Figure 1. *w*-relativisers diachronically, by social class.

In Figure 1, we plotted the proportion of *w*-relativisers across social class and time. As becomes clear from Figure 1, there is a strong social pattern with respect to the spread of *w*-relativisation. In the seventeenth century, *w*-forms constitute less than 20% in the lower class, whereas the lower-middle, upper-middle and upper class scores are 41, 45, and 51%, respectively. In the eighteenth century, the lower class is at this level of slightly over 40%, while the lower-middle class has risen to some 57%, and the upper-middle and upper classes approximate three-quarters *w*-forms of the total sum of relativisers used. Although the absolute numbers in the LC may be fairly low (Table 2), the established pattern nevertheless fits in with those of the other social classes, where numbers are higher. The incoming variant is most enthusiastically adopted by the upper-middle and upper classes, and is less prominent in the lower-middle and especially in the lower classes, where in the 1770s/1780s about 40% of the relative clauses are still
formed with a $d$-relativiser. In sum, the change is led by the upper-middle and upper classes and the change from $d$- to $w$-relativisation is clearly a change from above.

It should be noted that the LC score of 43% $w$-relativisers in the eighteenth century, corresponds to the seventeenth-century scores in the LMC and the UMC, which were 41% and 45%, respectively. In Chapter 4, analysing variation and change in epistolary formulae, we noted that female writers were often one social class behind male writers with respect to new writing practices, with, for example, UMC women scoring averages similar to those of LMC men. We referred to this phenomenon as *incremental writing experience* (Chapter 4, Section 4.3). We seem to find a similar effect in the case of $w$-relativisation, but here this phenomenon relates to the variables of time and social class, with eighteenth-century LC scores being similar to seventeenth-century LMC and UMC scores. Broadly speaking, we could say that LMC, UMC and UC writers are somewhere between 40 and 50 per cent in the seventeenth century, and have all risen to higher scores in the eighteenth century. With just over 40%, LC writers get into only the 40 to 50 per cent slot in the eighteenth century, while the LMC is lingering just below 60% and the UMC and UC are ahead with about 75%. We will take up this theme in the next section.

The next question concerns possible gender patterns influencing the distribution of $d$- and $w$-relativisers in the seventeenth and eighteenth centuries. Table 3 presents the absolute numbers of $d$- and $w$-forms across gender in both subcorpora.

Table 3. Relativisation in the seventeenth and eighteenth centuries, by gender.

<table>
<thead>
<tr>
<th></th>
<th>17th century</th>
<th></th>
<th>18th century</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$d$-relativiser</td>
<td>$w$-relativiser</td>
<td>Total</td>
<td>$d$-relativiser</td>
</tr>
<tr>
<td>M</td>
<td>94</td>
<td>96</td>
<td>190</td>
<td>106</td>
</tr>
<tr>
<td>F</td>
<td>56</td>
<td>23</td>
<td>79</td>
<td>63</td>
</tr>
<tr>
<td>Total</td>
<td>150</td>
<td>119</td>
<td>269</td>
<td>169</td>
</tr>
</tbody>
</table>

Figure 2 gives the proportion of $w$-relativisers across gender and time in both subcorpora, based upon the results in Table 3. Figure 2 shows that both in letters written by men and by women the frequency of $w$-relativisers rises in the eighteenth century relative to the seventeenth-century results. There is, however, in both periods a clear gender difference, in that men produce more $w$-forms than women. Whereas the proportion of $w$-forms in letters by men changes from approximately 50% in the seventeenth-century subcorpus to about 75% in the eighteenth century, the proportion in letters by women rises from some 30% to almost 60%. This difference is in line with earlier results, where we introduced the notion of incremental writing experience.

Thus we established both social class and gender differences in the use of $d$- and $w$-forms. We therefore cross-tabulated the absolute frequencies of $d$- and $w$-relativisation with both social class and gender in Table 4. Note that breaking down the results
in this way drastically diminishes the scores in a number of cells. Particularly the results for LC women in the eighteenth-century subcorpus are low: the only two letters in the subcorpus contain no more than two adverbial relative clauses.

Building on Table 4, we plotted the proportion of the incoming variant, that is of \textit{w}-forms, in Figures 3 and 4. Figure 3 gives the seventeenth-century results, Figure 4 the eighteenth-century results.

**Table 4.** Relativisation in the seventeenth and eighteenth centuries, by social class and gender.

<table>
<thead>
<tr>
<th></th>
<th>17th century</th>
<th></th>
<th>18th century</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>\textit{d}-relativiser</td>
<td>\textit{w}-relativiser</td>
<td>Total</td>
<td>\textit{d}-relativiser</td>
</tr>
<tr>
<td>LC-M</td>
<td>5</td>
<td>2</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>LC-F</td>
<td>5</td>
<td>–</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>LMC-M</td>
<td>8</td>
<td>13</td>
<td>49</td>
<td>21</td>
</tr>
<tr>
<td>LMC-F</td>
<td>21</td>
<td>7</td>
<td>28</td>
<td>5</td>
</tr>
<tr>
<td>UMC-M</td>
<td>68</td>
<td>66</td>
<td>163</td>
<td>21</td>
</tr>
<tr>
<td>UMC-F</td>
<td>21</td>
<td>8</td>
<td>48</td>
<td>23</td>
</tr>
<tr>
<td>UC-M</td>
<td>13</td>
<td>15</td>
<td>45</td>
<td>58</td>
</tr>
<tr>
<td>UC-F</td>
<td>9</td>
<td>8</td>
<td>17</td>
<td>33</td>
</tr>
<tr>
<td>Total</td>
<td>150</td>
<td>119</td>
<td>269</td>
<td>169</td>
</tr>
</tbody>
</table>
The black columns in Figure 3 represent the total scores per social class and are identical to the seventeenth-century columns in Figure 1 above. Figure 3 shows that the light grey columns representing male usage are consistently higher than the dark grey columns representing female usage. This means that within each social class a gender effect is established, in that more men than women use the incoming variant. Focusing on male writers, we see that the LC scores are just below 30\% \textit{w}-relativisers, whereas the LMC, UMC and UC scores are about or even over 50\%. Female writers, on the other hand, only reach this level in the UC, while the LMC and UMC scores lie approximately at the level of LC men, that is, just below 30\%. Among LC women, we find no adverbial relative clauses with a \textit{w}-form, but only five \textit{d}-forms instead. Again, we find the incremental writing experience established for epistolary formulae in Chapter 4. Men are ahead of women in adopting the new form.

Turning to the eighteenth-century distribution as visualised in Figure 4, similar patterns arise. The black total columns correspond to the results presented in Figure 1. Again, we find gender effects within the social classes, in particular in the UMC and the UC, and also in the LC, where the score for women is, however, based upon only two letters, as explained above. Male writers now start off at around 50\% in the LC, and then rise to around 80\% \textit{w}-relativisation in the UMC and the UC. Female writers remain between 50 and 60\% in the LMC and the UMC, with only a slightly higher score in the UC. The scores for women in the LMC and the UMC and UC, therefore, correspond to those of men in the LC and the LMC, again suggesting the incremental nature of writing experience mentioned before.
The hypotheses formulated in Section 3 were corroborated by the results. Both the seventeenth and the eighteenth-century results strongly suggest that the change from *d-* to *w*-relativisation was a change from above. First we found clear social patterns, with UC writers using the incoming variant at the highest rate, and LC writers remaining below 50% even at the end of the eighteenth century. Then we found equally clear gender patterns, with more women than men retaining the old form, and thus more men than women adopting the new one. The results therefore confirm our hypotheses concerning possible diachronic, social and gender patterns. It should be noted that these results tie in with research on English, in which the rise of *wh*-forms is also frequently assumed to have been a change from above (Rissanen 1999: 295), as well as with an earlier study of Dutch diaries (Rutten 2010), in which the one supposedly least-experienced writer was shown to produce the highest number of *d*-relativisers. In Section 3, we also related our hypotheses about variation and change in the relative clause to the notion of writing experience, as we considered the change from *d*- to *w*-relativisation as an example of syntacticisation. We will take up this theme in the next section.

5. Relativisation and incremental writing experience

Against the background of the syntacticising grammar of *written* Dutch, we expected that people who were more involved in the written culture would adopt the incoming
$w$-variant at the highest rate. In other words, we expected the distribution of the variants to be related to writing experience. It should be recalled from Chapter 4 that we operationalised writing experience using expectations about the involvement in the written culture founded upon social class and gender differences. In general, the upper and upper-middle classes were better educated, and written varieties of the language were more important in their daily lives than in the lives of people from the lower and lower-middle classes. A similar difference between being more and less connected to the written culture applies to men and women. Our case studies on formulaic language confirmed our ideas about writing experience (Chapters 4 and 5). We interpreted formulaic language as a safe option for less-experienced writers, and we demonstrated that there were clear social and gender patterns in the use of epistolary formulae, indicating that the use of formulae was connected to writing experience. We also found evidence of incremental writing experience. Frequently, for example, scores by UMC women turned out to be similar to those by LMC men, or scores by LMC women equalled those of LC men. In such cases, men appeared to be one social class ahead of women with respect to the decrease in the use of formulaic language.

The results reported in the previous section revealed that there was a similar social and gender distribution in the use of $w$-relativisers, and we have already alluded to incremental writing experience in Section 4.3. Figure 5 combines the seventeenth- and eighteenth-century results for male and female writers per social class, and presents the results in accordance with the proportion of $w$-forms per social group.

![Figure 5. $w$-relativisers in the seventeenth- and eighteenth centuries, by social class and gender.](image-url)
Figure 5 suggests that there are a number of different cut-off points. At the level of zero \(w\)-relativisers, we find both seventeenth- and eighteenth-century LC women. The next level is 25 to 30% \(w\)-forms, attested for seventeenth-century LC men. Seventeenth-century women reach this level in the LMC and the UMC. The third level is 50 to 65% \(w\)-relativisers. Here we find LMC, UMC and UC men in the seventeenth century. Of the females, it is only UC women who reach this level in the seventeenth century. In the eighteenth century, LC men have risen to 50% and LMC men are still in this range of 50 to 65% of the new variant, as are UC women. Women from the LMC and the UMC have also progressed to this level. The fourth level of approximately 80% new \(w\)-forms is reached by the male scores for the UMC and the UC in the eighteenth century. Incremental writing experience thus manifests itself at both intermediate levels: seventeenth-century middle-class women are at the same 25–30% level as lower-class men and eighteenth-century middle-class and upper-class women are at the same 50–65% level as lower-class (and LMC) men. Both extreme scores (zero and approximately 80%) are also gender-specific. Apart from the synchronic evidence of incremental writing experience, Figure 5 also indicates diachronic differences. This is most obvious for the eighteenth-century LMC and UMC women, whose scores are at the same level as those of seventeenth-century men from the LMC, UMC and UC, and of UC women. Diachronically, eighteenth-century women thus often write in the same way as seventeenth-century men. We also find this kind of diachronic evidence within the gender scores. The eighteenth-century score of LC men corresponds to the seventeenth-century scores of LMC, UMC and UC men. The eighteenth-century scores of LMC and UMC women correspond to the seventeenth-century score of UC women. As we said before, these patterns point clearly to the incremental nature of writing experience.

The case of relativisation thus allows us to extend the idea of writing experience to other grammatical domains, which are less obviously related to the influence of writing, literacy and occupation. Unlike epistolary formulae, relative clauses are not prototypical elements of a written text type, and they are certainly not restricted to written language. But there is variation in the formation of relative clauses, and this variation appears to be linked to writing experience. Those who write the most, use the more written-like \(w\)-forms, those who have less routine in writing, prefer the \(d\)-forms. We conclude that the textualisation of society and the increasing literacy rates have had an undeniable influence on the language. In other words, the written channel influences the grammar. We witnessed the decrease in local spellings in Chapter 2, indicating supraregional convergence, which implies deviation from the spoken language. Here, we also find such a deviation from the spoken language towards more typically written language, though in the field of morphosyntax.

6. Relativisation and formulaic language

Formulaic language stands out as a hallmark of the language found in our letters. Moreover, epistolary formulae are similar to constructions in the sense of construction
Table 5. The syntax-lexicon continuum in construction grammar (Croft & Cruse 2004: 255).

<table>
<thead>
<tr>
<th>Construction type</th>
<th>Traditional name</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex and (mostly) schematic</td>
<td>Syntax</td>
<td>[SBJ be-TNS VERB-PP by OBL]</td>
</tr>
<tr>
<td>Complex, substantive verb</td>
<td>Subcategorisation frame</td>
<td>[SBJ consume OBJ]</td>
</tr>
<tr>
<td>Complex and (mostly) substantive</td>
<td>Idiom</td>
<td>[kick-TNS the bucket]</td>
</tr>
<tr>
<td>Complex but bound</td>
<td>Morphology</td>
<td>[NOUN-s], [VERB-TNS]</td>
</tr>
<tr>
<td>Atomic and schematic</td>
<td>Syntactic category</td>
<td>[DEM], [ADJ]</td>
</tr>
<tr>
<td>Atomic and substantive</td>
<td>Word/lexicon</td>
<td>[this], [green]</td>
</tr>
</tbody>
</table>

grammar (e.g. Croft 2001; Goldberg 1995, 2006). This approach considers the entire language system made up of constructions as form/meaning pairings such as illustrated in the well-known Table 5, consisting of almost entirely schematic constructions on the one hand to fully lexicalised constructions on the other hand.

If the language system indeed consists of constructions such as in Table 5, it would follow that language change proceeds through constructions. As stated above (Section 2), Rutten (2010) argued that w-relativisers first occur in specific constructions, and are then generalised to a more abstract level. The fact that language change proceeds through different contextually-bound types is often referred to as lexical diffusion. Focusing on morphosyntactic constructions, we could term the process constructional diffusion. The constructions discussed in Rutten (2010) mainly revolved around certain verb types, such as verbs of cognition and communication (weten ‘know’, zien ‘see’), and verbs of movement (gaan ‘go’, rijden ‘ride’). Here, in the context of our letter corpora, we take another yet similar view. The most prominent constructions or ‘prefabs’ found in our seventeenth- and eighteenth-century letters are epistolary formulae (see Chapters 3–5). Epistolary formulae are what construction grammar calls complex constructions with (mostly) substantive elements. An example is the idiomatic expression [kick-TNS the bucket] in which TNS is the element open to variation. The question arises now whether the change from d to w proceeded through specific epistolary formulae, which are mostly substantive vis-à-vis the lexical material, but open to morphosyntactic variation. In this section, therefore, we discuss formulae in which the relative (pronominal) adverb appears, and we will see how these formulae behave with regard to the change from d- to w-relativisation.

6.1 Epistolary formulae

Since the seventeenth-century letters comprise more formulae than the eighteenth-century letters, as discussed in Chapter 4, we focus here on the seventeenth-century letters used for the present chapter. In these letters, formulae with a relative (pronominal) adverb occur and, interestingly, they appear to differ in their adoption of the incoming w-form.
The first formula, which mainly appears in addresses, identifies the captain of the ship to which the letters are sent. Example (10) has a $d$-form; (11) has a $w$-form. The formula is presented in boldface.

(10) desen bryef aen den eersammen man ijan wijnlemse luitenant op het schep this letter to the honourable man IJan Wijllemse lieutenant on the ship de spijegel daer op komder menheer menheer fijes amarael de ruijter De Spijegel rel on commands mister mister Vice Admiral De Ruijter ‘This letter to the honourable man IJan Wijllemse, lieutenant on the ship De Spijegel, on which the Vice Admiral Mr De Ruijter commands’

(11) Aen reijnier witte Chirurgijn op ’t slands schip de spiegel, waer op to Reijnier Witte surgeon on the country’s ship De Spiegel rel on Commandeert den hr viceadmirael michiel de ruijter commands the Mr Vice Admiral Michiel de Ruijter ‘To Reijnier Witte, surgeon on the country’s ship De Spiegel, on which the ViceAdmiral Mr Michiel de Ruijter commands’

The seventeenth-century autograph subcorpus contains 25 tokens of this formula. We compare the use of $d$- and $w$-relativisers in this formula to the overall frequency of $d$- and $w$-forms in the seventeenth-century autograph subcorpus. Since most tokens of the formula appear in the UMC (17 out of 25), due to the relatively high number of UMC letters in general, we also compare the distribution of $d$ and $w$ in this formula with the overall numbers of $d$ and $w$ in the UMC. See Table 6 for the results. Since numbers are low, we refrain from discussing possible differences linked to social class and gender.

Table 6. Frequency of daer/waer op commandeert ‘on which commands’ compared to the seventeenth-century autographs and to the UMC frequencies.

<table>
<thead>
<tr>
<th>17th-century</th>
<th>$d$-relativiser</th>
<th>$w$-relativiser</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>autograph subcorpus</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Total relative clauses</td>
<td>150</td>
<td>56</td>
<td>119</td>
</tr>
<tr>
<td>Of which in UMC</td>
<td>90</td>
<td>55</td>
<td>74</td>
</tr>
<tr>
<td>Formula daer/waer op commandeert Total</td>
<td>18</td>
<td>72</td>
<td>7</td>
</tr>
<tr>
<td>Of which in UMC</td>
<td>11</td>
<td>65</td>
<td>6</td>
</tr>
</tbody>
</table>

As can be seen in Table 6, the total autograph subcorpus as well as the UMC part contain about 55% $d$-relativisers. The formula daer/waer op commandeert generates more $d$-forms (65 to 72%) and thus appears to behave slightly more conservatively. Behaviour that is so much more conservative may not come as a surprise. As argued in Chapter 4, formulae were most probably memorised as a whole, and inserted into discourse as prefabricated strings of words. Not all formulae, however, behave conservatively, as we will show below.
We consider a second formula, which also behaves more conservatively than would be expected on the basis of the overall distribution of $d$- and $w$-forms. This formula can be divided into two subtypes which both express the subject’s emotional state. In (12) and (13), we present the schematic structures of both subtypes with the attested lexicalisations and their frequencies below. Both formulae begin with the relativiser and both contain a preposition, which is mostly in front of the verb, but occasionally appears next to the relativiser. The subject, if expressed, is in most instances pronominal.

The formula in (12) expresses the emotional state by an adjective, the formula in (13) by the verb. The verbs are given as infinitives.

(12) \text{rel} \ (\text{prep}) \ (\text{subj}) \ \text{adj}_{\text{EMO}} \ \text{prep} \ \text{V}

\begin{align*}
\text{daar} & \quad \text{blij} \ ‘\text{happy}’ \ (6) & \text{om} & \quad \text{‘about}’ \ (9) & \text{zijn} & \quad \text{‘are}’ \ (16) \\
\text{waar} & \quad \text{bedroefd} \ ‘\text{sad}’ \ (3) & \text{over} & \quad \text{‘about}’ \ (4) \\
& \quad \text{moeilijk} \ ‘\text{difficult}’ \ (2) & \text{mee} & \quad \text{‘with}’ \ (1) \\
& \quad \text{verwonderd} \ ‘\text{surprised}’ \ (2) & \\
& \quad \text{droevig} \ ‘\text{sad}’ \ (1) & \\
& \quad \text{ongerust} \ ‘\text{worried}’ \ (1) & \\
& \quad \text{verwonderd} \ ‘\text{dismayed}’ \ (1) & \\
\end{align*}

(13) \text{rel} \ (\text{prep}) \ (\text{subj}) \ \text{prep} \ \text{V}_{\text{EMO}}

\begin{align*}
\text{daar} & \quad \text{naar} \ ‘\text{to}’ \ (3) & \text{verlangen} & \quad ‘\text{long}’ \ (3) \\
\text{waar} & \quad \text{om} \ ‘\text{about}’ \ (1) & \text{zuchten} & \quad ‘\text{sigh, yearn}’ \ (1) \\
& \quad \text{in} \ ‘\text{in}’ \ (1) & \text{verwonderen} & \quad ‘\text{be surprised}’ \ (1) & \\
\end{align*}

A protoypical example of (12) is given in (14).

(14) \textit{daer ick seer blij om ben}  \\
\quad \text{‘about which I am very happy’}

The second subtype’s main variant contains the verb \textit{verlangen} ‘long’. Both other verbs (\textit{zuchten} ‘sigh, yearn’ and \textit{verwonderen} ‘be surprised’) are attested only once, in a similar syntactic pattern though, and both express emotions also found in the other formula, viz. sorrow and surprise. A protoypical example is given in (15).

(15) \textit{daer jck soo naer verlange}  \\
\quad \text{‘which I so much long for’}

Both formulae often follow statements on the writer’s health or confirmations that the addressee’s health has been successfully communicated, or the wish that the writer and addressee will be in touch again very soon. In other words, they often follow health formulae and contact formulae (see Chapter 3), thereby reinforcing the formulaicity of the expressions in question. We decided to combine the two subtypes in (12) and (13),

13. Zero subjects are common in the letters in our corpora.
which resulted in 21 tokens of these Emo formulae signalling emotional states. In Table 7, we compare the Emo formulae with the overall frequencies of the seventeenth-century autograph subcorpus and with the UMC frequencies, since most examples stem from UMC. As the numbers are low, we again refrain from discussing social and gender differences. Table 7 shows that the Emo formulae generate a relatively large number of $d$-forms (86 to 90%), in both the total autograph subcorpus and the UMC. The formulae behave conservatively with regard to the incoming $w$-forms.\footnote{In Rutten (2010: 22), a similar CAUSA-EMO construction was identified in diaries, which construction turned out to be one of the forerunners in the adoption of the new $w$-forms. In our letter corpus, on the contrary, EMO constructions prefer the conservative $d$-forms. This difference might be explained as a case of social variation: Rutten (2010) was based on diaries and journals mainly linked to upper and upper-middle classes whereas the letters used in the present chapter stem from various social classes, lower and lower-middle classes included.}

Apart from formulae behaving conservatively vis-à-vis the incoming $w$-relativisers, there are also formulae that behave progressively, in that they take up the incoming variant at a faster pace than expected on the basis of the average frequencies. One such formula we call $daer/waer$ Understand, whereby writers, after having acknowledged the receipt of a letter, describe what they have read in that letter, or what they understand that the addressee had written. Examples with $d$ and $w$ are in (16) and (17), respectively.

\begin{verbatim}
(16) Gonstige vrient ick hebbe ul aengenamen wel ontfang \textit{daer uit dat}
Kind friend I have your pleasant well received REL from that
\textit{ick versta dat ul getruit is}
I understand that you married
‘Kind friend, I have well received your pleasant [letter], from which I understand that you have married’

(17) soo ijst dat ick naer datto van dien een houder van datto uit\textit{t Cap}\
o is-it that I after date of that an older of date from captain
\end{verbatim}

\footnote{In Rutten (2010: 22), a similar CAUSA-EMO construction was identified in diaries, which construction turned out to be one of the forerunners in the adoption of the new $w$-forms. In our letter corpus, on the contrary, EMO constructions prefer the conservative $d$-forms. This difference might be explained as a case of social variation: Rutten (2010) was based on diaries and journals mainly linked to upper and upper-middle classes whereas the letters used in the present chapter stem from various social classes, lower and lower-middle classes included.}

### Table 7. Frequency of $daer/waer$ Emo compared to the seventeenth-century autographs and to the UMC frequencies.

<table>
<thead>
<tr>
<th>17th-century autograph subcorpus</th>
<th>$d$-relativiser</th>
<th>$w$-relativiser</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Total relative clauses</td>
<td>150</td>
<td>56</td>
<td>119</td>
</tr>
<tr>
<td>Of which in UMC</td>
<td>90</td>
<td>55</td>
<td>74</td>
</tr>
<tr>
<td>Formula $daer/waer$ Emo</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>19</td>
<td>90</td>
<td>2</td>
</tr>
<tr>
<td>Of which in UMC</td>
<td>12</td>
<td>86</td>
<td>2</td>
</tr>
</tbody>
</table>
Chapter 8. Variation and change in the relative clause

Tange hebbe ontfangen waer uijt verstaen ue grootelijcx verwondert
Tange have received rel. from understood you greatly surprised
zijt ick soo weijnich rettour ben zendende
are I so little return am sending

‘So it is [the case] that after the date of that letter I received a [letter] of an older date through captain Tange, from which I have understood that you are greatly surprised that I am returning so little’

The formula is usually located in the introductory part of letters, following the greeting, and often introducing the health formula. In the seventeenth-century autograph subcorpus, we identified 17 tokens of this daer/waer UNDERSTAND formula. Table 8 compares the frequency of the formula to the overall frequency of d- and w-relativisers and to the results for UMC, where most tokens are found.

Table 8. Frequency of daer/waer UNDERSTAND compared to the seventeenth-century autographs and to the UMC frequencies

<table>
<thead>
<tr>
<th></th>
<th>17th-century d-relativiser</th>
<th>17th-century w-relativiser</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>autograph subcorpus</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>150</td>
<td>56</td>
<td>119</td>
</tr>
<tr>
<td>Of which in UMC</td>
<td>90</td>
<td>55</td>
<td>74</td>
</tr>
<tr>
<td>Formula daer/waer UNDERSTAND</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4</td>
<td>24</td>
<td>13</td>
</tr>
<tr>
<td>Of which in UMC</td>
<td>2</td>
<td>14</td>
<td>12</td>
</tr>
</tbody>
</table>

As can be seen in Table 8, this formula generates about 76 to 86% w-forms, whereas the average distribution in the total autograph subcorpus as well as in the UMC part is approximately 55% d as opposed to 45% w. The formula behaves fairly progressively. We will come back to this below and in Section 7. First, we discuss another progressive formula.

The next formula is introduced by a relativiser, followed by a verb phrase intended to round off the conversation, such as eyndigende ‘ending’ or blijve ‘remain’. It can be paraphrased as with which I end/finish/round off this letter, or with which I remain. There are only eight tokens in the seventeenth-century autograph subcorpus. All tokens contain a w-relativiser; no d-forms occur. The formula appears towards the end of letters, rounding off what has been said so far and marking the transition to the closing. We will call it the waer END formula of which (18) and (19) are examples.

(18) Zal hem wel doen betaelen waermede Blijve met haest Waerde Moeije
Shall him well do pay rel.-with remain with hurry beloved aunt
uedw:d: en Neef Alexander Batij
your obedient servant and nephew Alexander Batij
‘[I] shall make him pay. With which I remain, in a hurry, beloved aunt, your obedient servant and nephew Alexander Batij’
(19) *ik ben klerk ter secretarie alhier waer mede eyndigende*, [...] I am clerk to the secretarial office here with ending *na hertelike groetenisse en wensinge alles goeds* [...] Nicolaes Pyl after hearty greeting and wishes all good Nicolaes Pyl ‘I am clerk to the secretarial office here. With which I end [this letter], [...] after kind regards and wishing [you] all the best [...] Nicolaes Pyl’

All eight tokens are from the UMC and the UC, and all are found in letters by male writers. It should be recalled from Figure 3 in Section 4.2 that male writers scored about 50% *d*-forms and 50% *w*-forms in the UMC as well as in the UC in the seventeenth century. In Table 9, we compare the frequency of the formula with the overall numbers of the seventeenth-century autograph subcorpus and with the numbers for the UC and the UMC. Note that we incorporate the UC here as well, because numbers would otherwise be even lower. Though numbers are low, it is still striking that all tokens carry the *w*-form. We will return to this below and in Section 7.

Table 9. Frequency of *waer END* compared to the seventeenth-century autographs and to the U(M)C frequencies.

<table>
<thead>
<tr>
<th>17th-century autograph subcorpus</th>
<th><em>d</em>-relativiser</th>
<th><em>w</em>-relativiser</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Total</td>
<td>150</td>
<td>56</td>
<td>119</td>
</tr>
<tr>
<td>Of which in U(M)C</td>
<td>111</td>
<td>53</td>
<td>97</td>
</tr>
</tbody>
</table>

Table 10 compares the frequency of the formula to the overall numbers and to the UMC numbers.

The last formula we would like to discuss here expresses gratitude to either the addressee or God. Examples are in (20–22).

(20) *daer myn liefste voor te dancken hebbe*

*REL my beloved for to thank have*

‘for which I have to thank my beloved’

(21) *daer wij godt almachtich niet cannen tot dancken*

*REL we God almighty not can to thank*

‘for which we are not able to thank God almighty’

(22) *waer uor ijck ul bedancke*

*REL for I you thank*

‘for which I thank you’

There are 19 tokens of this *daer/waer THANK* formula in the seventeenth-century autograph subcorpus, 14 of which occur in the UMC. Table 10 compares the frequency of the formula to the overall numbers and to the UMC numbers.
Table 10. Frequency of *daer/waer THANK* compared to the seventeenth-century autographs and to the UMC-frequencies.

<table>
<thead>
<tr>
<th>17th-century autograph subcorpus</th>
<th><em>d</em>-relativiser</th>
<th><em>w</em>-relativiser</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Total</td>
<td>150</td>
<td>56</td>
<td>119</td>
</tr>
<tr>
<td>Of which in UMC</td>
<td>90</td>
<td>55</td>
<td>74</td>
</tr>
<tr>
<td>Formula: <em>daer/waer THANK</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>53</td>
<td>9</td>
</tr>
<tr>
<td>Of which in UMC</td>
<td>8</td>
<td>57</td>
<td>6</td>
</tr>
</tbody>
</table>

It seems that this formula behaves completely in line with the overall tendency as described in the previous sections, viz. that about 55% of the relative clauses are formed with a *d*-relativiser, and about 45% with a *w*-relativiser.

Summing up, there are five epistolary formulae with a relative clause introduced by *daer* or *waer*. Two formulae (*daer/waer commandeert* and *daer/waer Emo*) behave conservatively. Considering that formulae are thought to be prefabricated units as a whole retrieved from memory, this does not come as a surprise. One formula (*daer/waer THANK*) appears to be in line with the general trend and two formulae, *daer/waer UNDERSTAND* and *waer END*, behave progressively. We do not think that this is evidence that these three formulae were not real formulae. On the contrary, the differences in the use of *d*- and *w*-forms between the formulae clearly demonstrate that the change spreads via specific formulae. We may visually represent the change from *d* to *w* as an arrow connecting 100% *d*-forms on the left, and 100% *w*-forms on the right. What we found in the present section is that formulae can be located at different points on this line, representing constructionally diffused differences in the distribution of *d*'s and *w*'s.

![Figure 6. Distribution of *d*- and *w*-relativisers in formulae.](image)

What is most surprising about these findings is that two formulae are clearly ahead of other formulae and other relative clauses in general, using many more *w*-forms than on average. We may wonder why these two formula, *daer/waer UNDERSTAND* and *waer END*, are the progressive ones. First of all, we would like to note that the latter formula is the prototypical example of a continuative relative clause, that is, a relative clause
that functions as a main clause. It has been argued before that continuative relative clauses have played an important role in the spread of \textit{w}-relativisers, in the history of English (Rissanen 1999: 293) as well as in Dutch (Rutten 2010; Rutten & van der Wal to appear). The present data seem to confirm this. We will come back to this in Section 7. Furthermore, the two formulae form the framework in which the informational and intersubjective part of the letter is placed. As noted in Chapter 3, the \textit{daer/waer UNDERSTAND} formula marks the transition from a reference to earlier contact – a recurrent topic in formulaic opening passages – to the intersubjective part. In other words, it marks the transition from text constitutive formulae to intersubjective formulae and/or non-formulaic intersubjective information. Consider example (16) again:

\begin{enumerate}
\item \textbf{(16) Gontige vrient ick hebbe ul aengenamen wel ontfang daer uijt dat}
\begin{enumerate}
\item Kind friend I have your pleasant well received \textit{rel} from that
\item \textit{ick versta dat ul getrout is}
\end{enumerate}
\begin{enumerate}
\item I understand that you married is
\end{enumerate}
\begin{enumerate}
\item ‘Kind friend, I have well received your pleasant [letter], from which I understand that you have married’
\end{enumerate}
\end{enumerate}

Example (16) is the opening passage of a letter, in which the writer acknowledges receipt of the addressee's letter and then focuses on new information: the addressee has married. This piece of new information is introduced by a formulaic relative clause, which does not add information to the preceding discourse but, on the contrary, starts off the discourse. Although the relativiser directly refers back to \textit{ul aengenamen} ‘your pleasant’ it does not thus introduce a restricted relative clause; if anything it introduces a continuative relative clause. External evidence for this is found in many similar examples where the formula appears as a coordinate clause, connected to the acknowledgement of receipt by the coordinator \textit{en} ‘and’, as in example (23).

\begin{enumerate}
\item \textbf{(23) soo laet ick u.l. weten als dat ick u.l. brief wel ontfangen ebbe ende daer}
\begin{enumerate}
\item so let I you know as that I your letter well received have and there
\item \textit{door verstaen als dat u.l. getrout zijt}
\end{enumerate}
\begin{enumerate}
\item through understood as that you married are
\end{enumerate}
\begin{enumerate}
\item ‘so I let you know that I have well received your letter and thereby/from it understood that you have married’
\end{enumerate}
\end{enumerate}

As in (16), the formula marks the transition to new, context-dependent and non-formulaic information, viz. that the addressee has married. While the relativiser in (16) refers back to a constituent in the preceding discourse, the formula actually points to information which is to come, and which is coded in the direct object of the verb form \textit{versta} ‘understand’. The intersubjective part of the letter, introduced by this formula, is then closed by the other formula under discussion that also prefers \textit{w}-relativisers. Consider example (18) again.
(18) Zal hem wel doen betaalen waarmede Blijve met haest Waerde Moeije
Shall him well do pay REL-with remain with hurry beloved aunt
UEDW:D: en Neef Alexander Batij
your obedient servant and nephew Alexander Batij
‘[I] shall make him pay. With which I remain, in a hurry, beloved aunt, your
obedient servant and nephew Alexander Batij’

When the writer finishes the intersubjective information, he may round off the inter-
subjective part by again using a continuative relative clause. Here, the antecedent of the
relative clause can be either the immediately preceding statement or the entire preced-
ing discourse, that is, the sum of all preceding intersubjective statements. The prag-
matic meaning of the formula is ‘this is it’ or ‘this is all I have to say’. What follows is,
in short, the writer’s name and his signature.

The intersubjective part of a letter may thus be framed by two w-preferring con-
tinuative relative clauses, which mark the transition from text constitution to intersub-
jectivity, and from intersubjectivity back to text constitution. These formulae, although
syntactically similar to other relative clauses, do not so much present informational
content relative to an antecedent, but rather indicate that a part of the discourse is end-
ing, and a new part is about to follow. Rather than referring back to what preceded,
they point forward to what is coming. The question rises as to what is so special about
continuative relative clauses, or about appositive relative clauses in general? Note that
other formulae, such as daer/waer EMO, which behaved conservatively, are also ap-
positive relative clauses. We will discuss these matters in Section 7.

6.2 Constructional diffusion

In this section, we have discussed the behaviour of formulae comprising relative clauses
vis-à-vis the change from d- to w-relativisation. Formulae were found to take up the
incoming w-forms at different paces. This confirms the idea that language change is
characterised by constructional diffusion. Questions arose about the special role of ap-
positive and particularly continuative relative clauses, and the degree of integration of
relative clauses into the main clause. These matters will be discussed in the next section.

7. Relativisation and clause integration

Rutten’s (2010) case study of seventeenth- and eighteenth-century diaries and journals
revealed that continuative relative clauses constitute not just a frequent construction
type, but also a type in which the incoming w-form is far more frequently employed
than the d-form. Similarly, Rissanen (1999: 293) remarked that “in the discussion of
the spread of the wh-forms [in the history of English], it has proved useful to distin-
guish a special type of non-restrictive clause called ‘continuative’”. He furthermore
stated that *wh*-forms are well established in the sixteenth century, “although *that* is still common in texts representing the oral mode of discourse […] There are, however, few unambiguous instances of *that* in continuative clauses” (Rissanen 1999: 293). Continuative relative clauses are thus not only considered carriers of the new form, but also the old form *that* is mainly found to occur in texts representing the oral mode of discourse. These observations are in line with what we have observed in Sections 4 and 5 above, viz. that *w*-forms are promoted through syntacticisation and textualisation. In Section 6, we identified two formulae that prefer the new variant, and frame the intersubjective discourse. Both formulae are continuative relative clauses, i.e. a particular type of appositive relative clauses. Against the background of Section 6 as well as Rissanen (1999), Rutten (2010), and Rutten & van der Wal (to appear), this calls for research into appositive relative clauses, and, in general, into the integration of relative clauses into main clauses.

7.1 Integration of relative clauses

For our study on the integration of relative clauses, we distinguish different categories of clause integration, combining syntactic and semantic clues. Syntactically, the relative clause’s degree of integration is determined by its position: embedded or final. Semantic integration depends on the relative clause being either restrictive or appositive. This leaves us with four options. We consider relative clauses as *expansions* of something that has already been mentioned (the antecedent). Adopting a linear approach to syntax (Auer 2009; Sinclair & Mauranen 2006), we first look at the surface-syntactic position where the relative clause is inserted. Two possible positions are attested: either immediately following the constituent it expands (24), or postponed to clause-final position (25). In the following examples, the antecedents and the relativisers are in boldface.

(24) *dese gaende met een cleen scheepje, waer* *op neeff Cornelis Meppelen gaet als assistent, sal alleen dienen*  
this going with a little ship+dim rel on cousin Cornelis Meppelen goes as assistant will only serve  
‘this [one, letter], sent with a little ship on which cousin Cornelis Meppelen works as an assistant, will only serve…’

(25) *dat zeij een poort hadden toe gesloeten waer* *doer dat de hollanders moesten pasceren*  
that they a gate had closed rel through that the Hollanders had to pass  
‘that they had closed a gate the Hollanders had to pass through’

In (24), the relative clause immediately follows the antecedent. Then, the main clause continues with the finite verb *sal* ‘will’, the subject of which is *dese* ‘this [one, letter]’. In
Lehmann’s (1984: 49) typology, this is an example of an embedded postnominal relative clause. In (25), the predicate *hadden toe gesloeten* ‘had closed’ of the subject *zeij* ‘they’ precedes the relative clause attached to *een poort* ‘a gate’. Following Lehmann (1984: 49), this is a relative clause in postposition. We will call examples such as (24) embedded, examples such as (25) final.

Semantically, we adopt the common distinction between restrictive and appositive relative clauses. The relative clause in (24) is restrictive: it would be pointless to state that the letter is sent with some little ship; what is decisive is the fact that it regards the ship on which Cornelis Meppelen works as an assistant. A syntactically similar construction is in (26), which however favours an appositive interpretation.

(26) *uE schrivens wegen mijn lossicheyt int vrije daer ul naer mijn oordeel al vrij wat gelooff in slaedt maeckt mijn gans geen onsteltenisse af*  
    ‘your writing about my moral laxity, to which you give quite some credit in my opinion, does not at all nullify my concern’

Example (24) presents an embedded restrictive relative clause, (25) contains a final restrictive relative clause, and (26) contains an embedded appositive relative clause. The formulae discussed in the previous section were all examples of the fourth option: final appositive relative clauses.

Building on this four-part categorisation of relative clauses, we examined the distribution of *d*- and *w*-relativisers over these categories and needed to add two more categories. First, free or headless relative clauses, as in the idiom in (27), cannot be classified in the present approach and will be regarded separately. There are 19 free relative clauses in the seventeenth-century autograph subcorpus. We also regarded grammaticalised relatives fulfilling an argumentative function separately, such as example (9), which is repeated here. There are 17 such cases in the seventeenth-century autograph subcorpus. Finally, 5 tokens were not categorised for lack of context, and thus left undecided.

(27) *daer men hovden daer vallen spander*  
    ‘you cannot make an omelette without breaking eggs’

(9) *god [...] dancken ende loeuen voor de genaede die heij aen ons beweijst*  
    ‘[we should] thank God and praise him for the mercy which He shows there where/even though we are such children of doom’
The research was carried out on the seventeenth-century autograph subcorpus, as this contains much more variation in the use of relativisers than the eighteenth-century subcorpus. It should be recalled from Table 1 in Section 4.1 that the distribution of $d$- and $w$-forms was 56% against 44% in the seventeenth century, whereas the eighteenth-century results provided already 72% $w$-forms. We allocated all 269 relative clauses in the seventeenth-century autograph subcorpus to the different types of relative clause as discussed in the foregoing. Table 11 shows the overall results of this categorisation.

Table 11. Distribution of $d$- and $w$-forms over categories of clause integration.

<table>
<thead>
<tr>
<th>Embedded</th>
<th>Final</th>
<th>Free relative</th>
<th>Argumentative</th>
<th>Undecided</th>
</tr>
</thead>
<tbody>
<tr>
<td>Restrictive</td>
<td>Appositive</td>
<td>Restrictive</td>
<td>Appositive</td>
<td>Restrictive</td>
</tr>
<tr>
<td>$d$</td>
<td>$w$</td>
<td>$d$</td>
<td>$w$</td>
<td>$d$</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td>17</td>
</tr>
<tr>
<td>Total</td>
<td>269</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Figure 7, we visualise the proportion of $d$- and $w$-forms for the main categories in Table 11 (embedded, final, free relative and argumentative), passing by the five tokens that were left undecided. Figure 7 clearly shows that embedded clauses highly prefer $d$-forms. There is only one example with a $w$-form, which is (25) above. The difference between restrictive and appositive does not seem to be of much importance here. Note that the majority of $d$-forms in embedded relative clauses, viz. 10 tokens, are from UC and UMC letters, and 8 tokens are produced by men, so that the $d$-preference in this category is not caused by an overrepresentation of, for instance, LC or LMC letters, or letters by women. In other words, though numbers are low, we do not think there is
any interference from social and/or gender effects. The one \textit{w}-form (25), however, was produced by a UC male, which could of course be linked to the writer's social class and/or gender. We conclude that the stronger syntactic integration of embedded clauses triggers the \textit{d}-forms. The adjacency of the antecedent brings the originally demonstrative function of the relativiser to the fore. As to the relatives with an argumentative function, it may not come as a surprise that these retained the historical \textit{d}-forms. Rutten (2010) argues that one reason why \textit{w}-forms took over the function of relativiser from the \textit{d}-forms is the functional overload of \textit{d}-forms. These functioned as relativisers, but also as demonstratives in assertive clauses, and they also grammaticalised into argumentative connectives.\footnote{Note that later on, \textit{w}-forms also grammaticalised into argumentative connectives. This development probably dates back to the eighteenth century – and probably the later part of the century, as the historical dictionary of Dutch, the \textit{WNT} only mentions nineteenth- and twentieth-century examples (\textit{WNT} s.v. \textit{waar VI}).}

Interestingly, free relatives prefer \textit{w}-forms in 53\% of all cases. This seems to recall van der Horst's claim that \textit{w}-forms first turn up in free relatives (see Section 2). Final relative clauses, however, also prefer \textit{w}-forms in 51\% of all cases. Whereas free relatives appear to behave progressively, final relative clauses are equally progressive. Moreover, the claim that the change began in free relative clauses and spread from there to other contexts, seems also improbable considering the low number of 19 free relatives compared to the very high number of 211 final relative clauses in this subcorpus.

Focusing on the difference between restrictives and appositives, we plotted the relevant results in Figure 8.
Figure 8 shows that the difference between restrictive and appositive does not influence the strong $d$-preference of embedded relative clauses. There is just one example with a $w$-form (25), which happens to be a restrictive relative clause. In final relative clauses, restrictives and appositives seem to behave differently. While the general trend for final relative clauses was about 50%–50% $d$- and $w$-forms (Figure 7), final restrictive relative clauses are slightly more conservative, preferring $d$-forms in 61% of all cases. Final appositives, on the other hand, prefer $w$-forms in 52% of all cases. So in final relative clauses, restrictives are more conservative than appositives, which use the incoming variant more often.\textsuperscript{16}

Summing up, we have distinguished different levels of clause integration. Syntactic integration as in embedded relative clauses triggers $d$-forms, thereby underlining the originally demonstrative function of the relativiser, which refers back to the adjacent antecedent. In the absence of syntactic integration, that is, in final relative clauses, semantically more integrated restrictive relative clauses prefer $d$-forms to a greater extent than semantically less integrated appositive relative clauses. The results presented in this section thus point to a link between clause integration and the distribution of $d$- and $w$-forms, more specifically, to a link between less integration and the incoming $w$-variants. Apart from the fairly infrequent free relative clauses, appositive relative clauses in final position behave most progressively. In the next section, we will discuss what the link between integration and relativisation strategies means, and how it relates to syntacticisation.

7.2 Integration and syntacticisation

What does it mean that final appositive relative clauses prefer $w$-relativisers to the greatest extent, and thus behave most progressively? From a cognitive perspective, the opposite could equally have been the case. As claimed by Bybee (e.g. 2006, 2010), highly frequent constructions become more entrenched. By consequence, we would expect these to behave conservatively in language change. Final appositive relative clauses are by far the most frequent relative construction in the seventeenth-century autograph subcorpus. No less than 183 out of 269 relative clauses (68%) are final appositive relative clauses (see Table 11). Of course, final appositive relative clauses may not be cognitively represented as one construction, and the conservative behaviour of some formulae (Section 6) clearly confirmed Bybee’s (2006) conserving effect of frequent strings.

\textsuperscript{16} In our view, there is no interference from social and/or gender effects, as for instance 68 (out of 87) $d$-forms in final appositive relative clauses are taken from UMC and UC letters, and 60 (out of 96) $w$-forms. So UC and UMC writers fairly equally distribute $d$’s and $w$’s which is what we would expect considering the distribution of $d$’s and $w$’s in general in these social classes (see Figure 1).
In our view, the syntacticisation approach taken in the present chapter can account for the fact that final appositive most frequently use the incoming variant. Following the discussion in the previous section, final appositive are the least integrated type of relative clauses. As stated above, we consider relative clauses as expansions of something that has already been mentioned (Auer 2009). According to Loock (2007), these least integrated expansions, that is, appositive relative clauses, come in three subtypes: continuative relative clauses, relevance relative clauses, and subjectivity relative clauses. The first subtype is mainly characterised by its discrepancy of form and function. Continuative relative clauses typically convey new information, which is commonly presented in a main clause, yet they have the form of a subclause. In conversation, they tend to have their own intonation contour. They belong to what are often called glue-ons or increments (Couper-Kuhlen & Ono 2007), that is, pieces of discourse which are syntactically linked to the immediately foregoing, but prosodically distinct. Continuative relative clauses create coherence with the preceding discourse by employing subordinating syntax where the information structure would canonically trigger a new main clause. The other two subtypes described by Loock (2007) are both used for detailing information in the main clause, which the speaker/writer deems necessary on second thoughts. Relevance appositive are a ‘repair strategy’ (Loock 2007: 346): what may have been not sufficiently determined in the main clause, is repaired by adding the appositive. Subjectivity appositive verbalise the speaker/hearer’s opinion, judgement, or a comment (Loock 2007: 353).

These different subtypes of appositive relative clauses share their relative lack of syntactic and semantic integration in the main clause. They continue discourse with something new, either with an entirely new topic – as in the case of continuative relative clauses –, or with a new perspective on the topic at hand – as in the case of relevance and subjectivity appositive. The main difference is the status of the information conveyed by the appositive. In continuative appositive the information is foregrounded and on the same level as the information in the preceding clause. In relevance and subjectivity appositive it is backrounded and secondary to the information in the main clause. The relative lack of integration of appositive is even greater when the relative clause is in final position. Adopting the incoming $w$-form, relative lack of integration of final appositive is repaired by the $w$, which marks subordination. In other words, syntacticisation of the relativisation system enables final appositive to become more explicitly integrated in the main clause, or in the case of continuative relative clauses, to become more explicitly connected to the preceding discourse.

The two rather conservative formulae discussed in Section 6 are relevance appositive ($daer/waer op commandeert$ ‘on which commands’) and subjectivity appositive ($daer/waer Emo$), respectively. The two most progressive formulae are clear cases of continuative relative clauses. As continuative relative clauses, they syntactically code discourse coherence and continuity, where no obvious semantic or informational coherence exists. As such, they create what Mithun (2008) calls dependency beyond the sentence. Consider example (19) again, which is repeated here.
(19) \( \textit{ik ben klerk ter secretarie alhier waer mede eyndigende, [...] na} \)
I am clerk to-the secretarial office here \textit{REL} with ending after \textit{hertelike groetenisse en wensinge alles goeds [...] Nicolaes Pyl}
hearty greeting and wishes all good Nicolaes Pyl
'I am clerk to the secretarial office here. With which I end [this letter], [...] after kind regards and wishing [you] all the best [...] Nicolaes Pyl'

The intersubjective and informational part of the letter is rounded off by the formula \textit{waer mede eyndigende} ‘with which I end’. There is no contextual clue as to why the writer considered the discourse sufficient once he wrote the word \textit{alhier} ‘here’. The writer informed the addressee of his occupation. A common continuation would have been information on how he got the job, what he thinks of it, what his daily activities are and the like. Immediately closing the letter with a greeting formula after \textit{alhier} would have created an awkward transition. Inserting a continuative relative clause indicating the act of closing prepares the reader for the actual closing of the letter. In other words, whereas the absence of the formula would have interrupted discourse continuity, the continuative appositive creates coherence. Relativisation here functions as a discourse coherence device. It syntactically codes a discourse function, viz. to keep the discourse going as we have shown in more detail in Rutten & van der Wal (to appear).

In sum, appositive relative clauses, and continuative relative clauses in particular adopt the incoming \textit{w}-forms at the fastest pace because they have the power of continuing discourse and creating coherence where discontinuity and incoherence threaten. They do so by syntactically coding discourse functions and as such, they represent a case of syntacticisation.

8. Conclusions

In this chapter, we have shown that the change from \textit{d}- to \textit{w}-relativisation was a change from above, led by those who were most involved in the written culture: UC and UMC men. We explained the distributional patterns found with reference to syntacticisation and writing experience: stronger marking of relative clauses is linked to being more involved in the written culture, as the written language is more in need of explicit syntactic marking. As in our studies on formulaic language (Chapter 4), we found evidence of incremental writing experience. Since formulaic language is so important in our letter corpora (Chapters 3–5), we focused on the distribution of \textit{d}- and \textit{w}-relativisers in epistolary formulae with relative clauses. Formulae adopt the incoming \textit{w}-relativisers at different paces, confirming the conclusion in Rutten (2010) that language change proceeds via constructions, and that we can therefore speak of constructional diffusion. While some formulae were conservative, preferring \textit{d}-forms, others were progressive, particularly continuative relative clauses. This finding led to an investigation of the integration of relative clauses. We found that, generally speaking, embedded
relative clauses were most conservative, and appositive relative clauses in final position most progressive. This we also interpreted within the syntacticisation framework: the relatively low integration of appositives into the main clause is repaired by adopting a syntactic strategy that explicitly codes a pragmatic function, viz. the function of keeping the discourse going.
Chapter 9

Apocope of final schwa

1. Introduction

One of the most salient changes in the history of Dutch is apocope of final schwa, a linguistic phenomenon that also occurred in English and to a lesser extent in German (Schönfeld/van Loey 1954: 109–110; van Bree 1987: 79–80; cf. Lass 1992: 79–81; Salmons 2012: 238). It is often considered a Modern Dutch development, though the first schwa-less forms date back to around 1300 (van Loon 1986: 97). This may be the reason that Marynissen (2009: 235) refers to apocope as one of the main phonological changes in Middle Dutch. Apocope of final schwa affected various word classes: nouns (e.g. Middle Dutch hase > Modern Dutch haas ‘hare’), adjectives (bose > boos ‘angry’), adverbs (lange > lang ‘long’), articles (ene > een ‘a’), pronouns (haerre > haar ‘her’), numerals (achte > acht ‘eight’), prepositions (ane > aan ‘on’), conjunctions (ende > en ‘and’) and also verbs. In the category of verbs, schwa apocope occurred in the first person singular of the present indicative (ic leve > ik leef ‘I live’), the first person singular of the past indicative of irregular weak verbs (ic brachte > ik bracht ‘I brought’), and inflected infinitives (te doene > te doen ‘to do’).

Although schwa apocope affected the lexicon to a large extent, various constraints also appear to apply. For instance, final schwa is often maintained after $d$ in lexemes such as schade ‘damage’ and hulde ‘tribute’ (Boutkan & Kossmann 1998). Another important, morphological constraint concerns the past indicative forms of regular weak verbs, where the schwa has a grammatical function, distinguishing the past from the present, for instance in hij werkte ‘he worked’, where *hij werkt would be similar to the present hij werkt ‘he works’. The schwa is also maintained in lexicalised phrases such as ten name van ‘in the name of’, in koelen bloede ‘in cold blood’ and zegge and schrijve ‘lit. (I) say and (I) write, precisely, no more than’ (van Loon 1986: 98) against the usual nouns naam ‘name’ and bloed ‘blood’, and the verbs zeg ‘say’ and schrijf ‘write’. Despite these and other constraints, apocope of final schwa is generally considered a primarily phonological change, with preceding $d$ constituting the most conservative context (Boutkan & Kossmann 1998).

By the seventeenth century, many dialects and particularly Holland dialects had a high proportion of schwa-less forms. The oldest schwa-less forms are dated around 1300 and stem from Holland, Utrecht, the south of Gelderland and Limburg, that is from the north-west, middle and south-east of the language area, from where apocope gradually spread to other areas in subsequent periods (van Loon 1986: 97). Apocope probably first affected polysyllabic nouns such as riddere ‘knight’, indicating the
Letters as Loot

phonological or rather prosodic nature of the phenomenon (Marynissen 2009: 237–239). Van Loey (1980: 55–56) notes that schwa-less forms in the first person singular of the present indicative appear in the fifteenth century, though earlier examples are found in specific contexts such as subject-verb inversion in which the clash of two vowels may promote apocope (e.g. *groetic* < *groete ic* ‘greet I’). At present, absence of apocope still marks the dialects of the south-west and of the east and north-east, which is clearly shown on the maps of the first person singular present indicative in the MAND (vol. II, maps 4.6.1.1 *ik breek* ‘I break’, 4.6.1.13 *ik leef* ‘I live’, and 4.6.1.23 ‘paradigm present singular *break, have, knock, live, get*’). Importantly, the province of Zeeland is at the crossroads of apocope and absence of apocope. In the south-west of the language area, final schwa is maintained in many dialects, while Holland to the north of Zeeland, and Brabant to the east are characterised by apocope.

While the beginnings of apocope of final schwa must be dated to the earliest Middle Dutch, it is in the first person singular present indicative in particular that variation existed well into the seventeenth century (Weijnen 1956: 38). In the language of Michiel de Ruyter (1607–1676), born and bred in Zeeland, schwa forms outnumber forms with apocope (Koelmans 1959: 125). But also outside Zeeland, in the works of a Holland-based author such as Joost van den Vondel (1587–1679), verbs both with and without final schwa appear (van Helten 1881: 38). Forms with final schwa dominate in the Authorized Version of the Bible (1637), though forms without schwa also occur, in particular in the case of subject-verb inversion (e.g. *loop ick* ‘walk I’, *heb ick* ‘have I’; Heinsius 1897: 70). Koelmans (2001: 68) notes that in De Ruyter’s language, inversion does not seem to influence the form of the finite verb.

Given this variation in first person singular present indicative forms in the seventeenth century, the present chapter will focus on these forms. In addition, the first person generates a considerable number of tokens in the language of letters, thus allowing for quantification. Since apocope is generally considered a phonological change, spreading from region to region, we will first discuss the variation in first person singular present indicative forms from this perspective (Section 3). In Sections 4 to 7, we will discuss the social variation found in the various regions under discussion, viz. North Holland, South Holland and Zeeland. In Section 8, we will examine whether specific contexts characteristic of the language of letters inhibit apocope. As mentioned above, lexicalised phrases may maintain forms with final schwa, even up to the present day. Since the letters in our corpora are characterised by formulaic language, such a conservative effect could be expected to have existed. Another distinguishing feature of the language of both letters and diaries is the frequent absence of subjects. We will also investigate whether this has any influence on the choice of variants. Before embarking upon final schwa apocope as a regionally diffused phonological change, we will first introduce the subcorpora used for this chapter in Section 2.
Chapter 9. Apocope of final schwa

2. The subcorpora

The eighteenth-century subcorpus that we used for the present chapter is the total eighteenth-century subcorpus, which consists of only autographs (cf. Chapter 1, Section 6). For our regional research in Sections 3 to 7, however, we focused on the four major regions of North Holland, Amsterdam, South Holland and Zeeland (cf. Chapter 1, Section 5), which leads to a main regions eighteenth-century subcorpus of 274 letters. With regard to the seventeenth-century data, in the previous Chapters 4, 6 and 8, we used only established autographs by writers whom we were able to assign to a social class index (viz. 219 letters). The seventeenth-century autograph subcorpus is in fact somewhat larger (viz. 260 letters; cf. Chapter 1, Section 6), as it also comprises a fair number of letters that we were unable to assign to a specific social class, but that nevertheless can be used for research on gender and age. For the present chapter we used almost the entire seventeenth-century private subcorpus including both the autograph subcorpus of 260 letters and, in addition, a large number of seventeenth-century letters that could be assigned to the four major regions of North Holland, Amsterdam, South Holland and Zeeland. This main regions seventeenth-century subcorpus thus comprises 454 letters, autographs as well as established non-autographs and letters of uncertain autograph status.1 Table 1 presents the details of the four subcorpora used for the present chapter. The largest seventeenth-century subcorpus (the main regions subcorpus) is used for Section 3 as well as for Section 8.1, while the seventeenth-century autograph subcorpus is used for Sections 4 to 7 and 8.2, in which social variation will be taken into account. The total eighteenth-century subcorpus (all autographs) was used for Sections 8.1 and 8.2; the main regions eighteenth-century subcorpus was used for Sections 3 to 7.

Table 1. Basic numbers of the subcorpora used for the present chapter.

<table>
<thead>
<tr>
<th>Period</th>
<th>Subcorpus</th>
<th>N letters</th>
<th>N words</th>
</tr>
</thead>
<tbody>
<tr>
<td>1660s-1670s</td>
<td>main regions</td>
<td>454</td>
<td>201,000</td>
</tr>
<tr>
<td>1660s-1670s</td>
<td>autograph</td>
<td>260</td>
<td>118,000</td>
</tr>
<tr>
<td>1780s</td>
<td>total</td>
<td>384</td>
<td>196,000</td>
</tr>
<tr>
<td>1780s</td>
<td>main regions</td>
<td>274</td>
<td>136,500</td>
</tr>
</tbody>
</table>

1. From the entire seventeenth-century private subcorpus of 549 letters (cf. Chapter 1, Section 6), letters that could not be assigned to a specific region and a number of letters from other regions such as Friesland and Flanders were excluded.
3. A regionally diffused phonological change

3.1 The broad picture by time and region

In this section, we discuss phonological and regional aspects of the sound change known as schwa apocope, focusing on the forms of the first person present indicative. We extracted the relevant forms from the subcorpora by searching for first person pronouns, notably ic, ick and ik ‘I’. This implies that we did not extract those forms where the first person pronoun is omitted. For the absence and presence of subject pronouns, and the effects of this variation on the form of the finite verb, see Section 8.2.

A number of verbs were excluded from the analysis. Apart from the irregular verb zijn ‘be’, which historically does not have forms with final schwa but forms such as ben and bin ‘am’, we also excluded monosyllabic verbs and modals. The monosyllabic verbs doen ‘do’, gaan ‘go’, slaan ‘hit’, staan ‘stand’ and zien ‘see’ as well as their derivatives and compounds such as verstaan ‘understand’ do not have final schwa in the first person singular either. The most common form of the first person singular present indicative of these verbs, from historical times up to present-day standard Dutch, is the so-called zero form, identical with the stem, viz. doe, ga, sla, sta, zie. In present-day dialects in Holland and Zeeland, zero forms are well attested, but so are forms with final n and final t, for example ik doen and ik doet ‘I do’ (see e.g. MAND, vol. II, map 4.6.1.22). These are also found in our corpora but will not be taken into account, as the variable we are interested in concerns the presence or absence of final schwa, which never occurs with these verbs. The modals kunnen ‘can’, moeten ‘must’, mogen ‘may’, willen ‘want’ and zullen ‘shall’ will not be taken into account either. Historically, these verbs did not have final schwa in the first person singular since they are preterito-presentia, that is present tense forms derived from (schwa-less) preterite forms. Technically, this does not apply to willen, but in the history of Dutch it behaves as the preterito-presentia (see e.g. WNT s.v. willen). In line with their origin, these verbs do not occur with final schwa in our corpora, apart from one or two isolated seventeenth-century tokens of ic wille ‘I want’. Since willen is a frequent verb, and since first person present indicative forms of willen with final schwa do not occur in the eighteenth-century subcorpus, we decided to disregard willen altogether. It should be noted that we did include weten, which is another preterito-presens, historically lacking final schwa in the first person singular. Forms with final schwa are not attested in the historical dictionaries of Dutch (ONW, VNMW, MNW, WNT s.v. ‘weten’), but in our subcorpora we do find analogical forms such as ik wete, both in the seventeenth and the eighteenth centuries.

The corpus search generated 4285 results, 2557 for the seventeenth century and 1728 for the eighteenth century (see Table 2 below), distributed over the different variants. The most frequent variant in both periods is the zero form, where, in other words, apocope of final schwa has taken place. Examples (1a, b) are from the seventeenth century, (1c, d) from the eighteenth century. The relevant first person present indicative forms are in boldface.
Chapter 9. Apocope of final schwa

(1) a. ick en twyffel niet of wij sulle alle jaeren wel bryeve van
I NEG doubt not if we shall all years well letters of
mekander hebbe
each other have
‘I do not doubt that we will have letters from each other each year’
b. ick stuér u met desen schepen 2 brieven
I send you with these ships two letters
‘I send you two letters with these ships’
c. In ’t Nieuws Aangaende de Famielje neem ik veel belang
in the news concerning the family take I much interest
‘I am very interested in news about the family’
d. ik wens u veel geluk En zeegen in het nuiwe jaar
I wish you much happiness and blessing in the new year
‘I wish you much happiness and blessings in the new year’

The second most frequent category of tokens is constituted by the historical forms with final schwa unaffected by apocope. Examples (2a, b) are from the seventeenth century, while (2c, d) are taken from the eighteenth-century subcorpus.

(2) a. Ick en twiffele niet of dat zijn maer al valsche tongen
I NEG doubt not if that are but all false tongues
‘I do not doubt that those are all scandalmongers’
b. vorders stiere ick vl twee kinneties harinc
further send I you two barrels+DIM herring
‘further, I send you two small barrels of herring’
c. neeme ik de vryheid uwelEdel deede te schrijven
take I the liberty your honour this to write
‘I take the liberty of writing this to your honour’
d. Jk wensche gaarn te weeten, of uwEd: mijn haan aan Zussie
I wish gladly to know if you my cock to sister+DIM
styntji bezorgt heeft
Styntji delivered have
‘I would like to know if you have delivered my cock to little sister Styntji’

The third set of tokens ends in -en. These are far less frequent than forms with apocope and forms with the historical schwa. The forms in -en maintain a schwa element with the addition of the ⟨n⟩ which was historically not present. This addition could be an effect of the reverse: in many forms such as infinitives (e.g. hebben ‘have’, hopen ‘hope’, schrijven ‘write’), plural present and past tense forms (e.g. wij hebben ‘we have’, wij had- den ‘we had’), and plural nouns (e.g. brieven ‘letters’, schepen ‘ships’), final n is dropped in large parts of the Dutch language area (e.g. MAND, vol. II, map 4.6.1.10 wij breken ‘we break’; MAND, vol. I, maps 1.6.1.1-1.6.1.5 beren ‘bears’, ogen ‘eyes’, taarten ‘cakes’,
Letters as Loot

ribben ‘ribs', bedden ‘beds’). Deletions of final n are well attested in the history of Dutch and in our corpora as well (van der Wal e.a. 2012). Contrary to final schwa, where writing practices evolved from schwa to zero in line with the sound change, there is a long writing tradition that realises final n despite its deletion in the spoken language. This continues up to the present day with standard spelling still prescribing final n. Since there was this long and firm writing tradition of realising final n where it was not pronounced, the reverse phenomenon of attaching a final n where it was absent historically may also occur (e.g. Koelmans 2001: 68, 197). Examples of this might be (3a), seventeenth-century beginnen (instead of beginne ‘I begin’) and (3b), eighteenth-century zieten (instead of zitte ‘I sit’), although we should not exclude the possibility that the -en suffix was actually pronounced, knowing that a few present-day dialects in the south-west of the Dutch language area show final en in first person singular present indicative forms (MAND, vol. 1, map 4.6.1.1 ik breek ‘I break’). Therefore, we consider these en-occurrences as a separate category, even though they are similar to the forms with final schwa in the sense that the historical schwa is maintained.

(3) a. want ick nu voortaan beginnen nae ul ende t’ Patria te for I now from now on begin to you and the fatherland to long
‘for now I begin to long for you and the fatherland from now on’

b. denkt dan Eens myn geliefde man in hoe een groot think+IMP then once my dear husband in how a great
verlangen dat ik zieten desire that I sit
‘please think, then, my dear husband, how strong the desire is that I have’

Finally, there is a group of what we will call other endings. In the seventeenth century, this group comprises forms in -t, and a handful of forms with either an apostrophe or a final n, and no schwa (4a-c). In the eighteenth century, the other endings are limited to final t (4d).

(4) a. ijk een geloft oock niet dat ghij een brief uan mijn uijt I NEG believe also not that you a letter from me through leest read
‘I also do not believe that you read through a letter from me’

2. These West- and French-Flemish dialects and Flemish dialects of the Dutch province of Zeeuws-Vlaanderen are spoken in areas in the south of and immediately to the south of Zeeland. Therefore an earlier extension of the -en pronunciation in other parts of Zeeland may be possible.
b. *zend’ ik u bij dese, 2 Boecken*
   send I you with this two books
   ‘I herewith send you two books’

c. *al soo ick verhopn dat het met ul mede soo is*
   as such I hope that it with you also so is
   ‘I hope that you are also like that [i.e. in good health]’

   d. *ik denkt tuis te weesen in de maant van meij*
   I think home to be in the month of May
   ‘I expect to be home in the month of May’

First person singular present indicatives with final *t* are well attested in present-day dialects, in South Holland in particular, but they occur also in other regions such as North Holland (*MAND*, vol. II, map 4.6.1.1; de Vogelaer 2008: 66). In our corpus, they are found in Zeeland as well. Forms with an apostrophe diacratically express apocope of final schwa: the apostrophe, in other words, signals the position historically occupied by final schwa. The tokens with final *n* may be slips of the pen for *en*, but as there are many present-day dialects with final *n* and without a schwa, especially in the southwest of the language area, in the south of Zeeland and in West and French Flemish (*MAND*, vol. II, map 4.6.1.1), we allocated these to the group of other endings. Note, in this respect, that example (4c) occurs in a letter written by someone originating from Zeeland.

Table 2 presents the overall results of the first person singular present indicative endings found in the seventeenth-century main regions subcorpus and in the eighteenth-century main regions subcorpus. As can be seen in Table 2, zero forms stand out in both periods. In the seventeenth century, they make up 59% of all tokens, rising to 73% in the eighteenth century. Forms with schwa decrease from (33+5=) 38% in the seventeenth century to (19+3=) 22% in the eighteenth century. The increase in zero forms and the concomitant decrease in forms with schwa clearly demonstrate the ongoing language change. The other endings make up a modest proportion in both periods, viz. 3% in the seventeenth century and 5% in the eighteenth century.

Since there is so much regional variation in the present-day dialects, and since our subcorpora comprise letters from various regions, we split the overall results according to region, focusing on the four major regions represented in the subcorpora, viz. North

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>17th century</strong></td>
<td>1,516</td>
<td>59</td>
<td>853</td>
<td>33</td>
<td>120</td>
<td>5</td>
<td>68</td>
<td>3</td>
<td>2,557</td>
<td>100</td>
</tr>
<tr>
<td><strong>18th century</strong></td>
<td>1,269</td>
<td>73</td>
<td>324</td>
<td>19</td>
<td>56</td>
<td>3</td>
<td>79</td>
<td>5</td>
<td>1,728</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 2. First person singular present indicative endings, overall results diachronically.
Table 3. First person singular present indicative endings by region, seventeenth century.

<table>
<thead>
<tr>
<th>Region</th>
<th>〈ø〉</th>
<th>〈e〉</th>
<th>〈en〉</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>North Holland</td>
<td>362</td>
<td>69</td>
<td>146</td>
<td>28</td>
<td>11</td>
</tr>
<tr>
<td>Amsterdam</td>
<td>701</td>
<td>72</td>
<td>191</td>
<td>20</td>
<td>48</td>
</tr>
<tr>
<td>South Holland</td>
<td>308</td>
<td>73</td>
<td>85</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>Zeeland</td>
<td>145</td>
<td>23</td>
<td>431</td>
<td>68</td>
<td>47</td>
</tr>
<tr>
<td>Total</td>
<td>1,516</td>
<td>59</td>
<td>853</td>
<td>33</td>
<td>120</td>
</tr>
</tbody>
</table>

Holland (excluding Amsterdam), Amsterdam, South Holland and Zeeland. Table 3 presents the seventeenth-century results. Recall from Table 2 that the overall proportion of zero forms was 59% in the seventeenth century. Table 3 shows that this proportion is actually higher, viz. approximately 70% in the three most northern regions, where simultaneously the proportion of forms with schwa is much lower than the 38% presented in Table 2. In North Holland, Amsterdam and South Holland, the proportion of forms with schwa varies from 23% to 30%. The opposite pattern is found in Zeeland, where schwa is maintained in (68+7=) 75% of all tokens, zero forms taking up no more than 23%. In sum, there is a clear regional difference, with the three northern regions behaving similarly, while Zeeland diverges from this pattern. The Zeeland results show that apocope of final schwa probably spread from the centre of the language area (Holland/Brabant) to the south-west, as is also suggested by the present-day dialect situation, where forms with schwa can still be found in parts of Zeeland.

Table 4 presents the eighteenth-century results, divided by region, and shows that the situation in Amsterdam and South Holland has not changed dramatically, that is in terms of the overall distribution of the variants. Zero forms appear in slightly over 70%, while the proportion of forms with schwa is just over 20%. In North Holland, the northernmost region in the subcorpus, the proportion of zero forms has risen to 84%, approaching what Nevalainen & Raumolin-Brunberg (2003: 55) consider a completed

Table 4. First person singular present indicative endings by region, eighteenth century.

<table>
<thead>
<tr>
<th>Region</th>
<th>〈ø〉</th>
<th>〈e〉</th>
<th>〈en〉</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>North Holland</td>
<td>121</td>
<td>84</td>
<td>18</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>Amsterdam</td>
<td>714</td>
<td>73</td>
<td>173</td>
<td>18</td>
<td>36</td>
</tr>
<tr>
<td>South Holland</td>
<td>291</td>
<td>71</td>
<td>90</td>
<td>22</td>
<td>8</td>
</tr>
<tr>
<td>Zeeland</td>
<td>143</td>
<td>71</td>
<td>43</td>
<td>21</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>1,269</td>
<td>73</td>
<td>324</td>
<td>19</td>
<td>56</td>
</tr>
</tbody>
</table>
change, that is when the incoming variant occupies over 85% of all tokens. Between 66 and 85%, the change is ‘nearing completion’. Interestingly, Zeeland is now completely in line with the other regions, in that zero forms are responsible for 71% of the tokens, while forms with schwa make up (21+5=) 26%. In other words, the change is also nearing completion in Zeeland. Despite the fact that the present-day dialect situation still shows a varied picture in Zeeland, with both schwa forms and zero forms, the written language of the eighteenth century patterns similarly to South Holland and Amsterdam.

3.2 Right context

As discussed by Booij (1995: 150–151) with regard to present-day standard Dutch, and as is well-known from many other languages, hiatus may be avoided in connected speech. Hiatus breakers are inserted consonantal elements such as glides (5a), but n (5b) and, in English, r (5c) may also be used as linking elements, while schwa deletion is another option (5d). Cf. the examples (5a), (5b), (5d) (Booij 1995: 150) and (5c) (Lass 1999: 114).

(5) a. Henk *haalde de vlo [v] uit het eten* 
   Henk removed the flea from the food
   ‘Henk removed the flea from the food’

   b. Henk *haalde [n] een vlo uit het eten* 
   Henk removed a flea from the food
   ‘Henk removed a flea from the food’

   c. *law [r] and order* 
   ‘law and order’

   d. *rod-ø aardbeien for rode aardbeien* 
   red strawberries

With regard to historical Dutch first person singular present indicative forms, it has been suggested that vocalic right contexts, as in the case of subject-verb inversion would trigger schwa apocope (Heinsius 1897: 70; van Loey 1980: 55–56). It may, therefore, be hypothesised that zero forms are promoted when the finite verb is followed by a word with an initial vowel. In other words, vocalic right contexts could be expected to be more progressive than consonantal right contexts. Similarly, en-endings could be interpreted as historical schwa forms with an additional hiatus-breaking n attached to them, in which case we would expect en-endings to prefer vocalic right contexts.

3. Nevalainen & Raumolin-Brunberg (2003: 54–55) use the following classification in terms of the proportion of incoming variants: 1 incipient, below 15%, 2 new and vigorous, between 15 and 35%, 3 mid-range, between 36 and 65%, 4 nearing completion, between 66 and 85%, 5 completed, over 85%.
In order to examine the possible influence of the right context, we investigated the interplay of first person endings and vocalic and consonantal right contexts. The seventeenth-century results are presented in Table 5. It should be noted that the total number of 2,529 tokens is lower than the 2,557 tokens in previous tables due to the fact that a number of finite verbs occurred at the end of a sentence or paragraph and thus lacked a vocalic or consonantal right context.

Table 5 shows that vocalic right contexts attract 65% zero forms, while consonantal right contexts elicit 53% zero forms. This means that the incoming schwa-less variant is more likely to occur when the following word begins with a vowel, and that there is a modest effect of the phonological right context. We will shortly turn to the eighteenth-century results and to the distribution across region, but we first consider Table 6, where we concentrate on the distribution of \(\langle e\rangle\) and \(\langle en\rangle\) in order to investigate whether forms with final schwa attract \(n\) as a linking element before vowels.

Table 6 clearly shows that the distribution of \(\langle e\rangle\) and \(\langle en\rangle\) does not vary in accordance with the phonological right context. This means that we will not consider final \(n\) in forms with \(\langle en\rangle\) as a linking element (see also Section 3.1 above on the interpretation of final \(n\)).

Interestingly, when examining the influence of vocalic and consonantal right contexts on the distribution of first person endings in the eighteenth century, a different picture emerges; see Table 7.
Table 7. First person singular present indicative endings by right context, eighteenth century.

<table>
<thead>
<tr>
<th>18th century</th>
<th>ø</th>
<th>⟨e⟩</th>
<th>⟨en⟩</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>vowel</td>
<td>672</td>
<td>75</td>
<td>181</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>900</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>consonant</td>
<td>586</td>
<td>73</td>
<td>136</td>
<td>17</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td>805</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1,258</td>
<td>74</td>
<td>317</td>
<td>19</td>
<td>53</td>
</tr>
<tr>
<td></td>
<td>1,705</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8. Schwa endings of the first person singular present indicative by right context, eighteenth century.

<table>
<thead>
<tr>
<th>18th century</th>
<th>⟨e⟩</th>
<th>⟨en⟩</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>vowel</td>
<td>181</td>
<td>89</td>
<td>203</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>consonant</td>
<td>136</td>
<td>81</td>
<td>167</td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>317</td>
<td>86</td>
<td>370</td>
</tr>
<tr>
<td></td>
<td>53</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>100</td>
</tr>
</tbody>
</table>

In the eighteenth century, any convincing influence of the phonological right context has completely disappeared. Both vocalic and consonantal right contexts trigger about 74% zero forms, fully in line with the results presented in Tables 2 and 4 above. The conclusion that final n is not a linking element is furthermore confirmed by a comparison of forms with ⟨e⟩ and ⟨en⟩; see Table 8. Contrary to Table 6 with the seventeenth-century results of this comparison, there now appears to be a small difference between vocalic and consonantal right contexts with regard to the distribution of ⟨e⟩ and ⟨en⟩. The difference is not as expected, however, as it becomes apparent that ⟨en⟩ is more frequent when preceding a consonant (19%) than when preceding a vowel (11%). The conclusion remains that final n should not be interpreted as a hiatus-breaking linking element.4

Since apocope of final schwa is first and foremost a regionally diffused phonological change, we also looked at the interplay of first person endings and right contexts in each of the four major regions. Figure 1 presents the seventeenth-century results. It becomes clear from Figure 1 that the influence between vocalic and consonantal right contexts established in Table 5 above varies across region. In North Holland, Amsterdam and South Holland, the incoming zero variant is always above the level of 60%. When followed by a vowel, the proportion of zero forms (the black columns)

---

4. We applied statistical tests only in cases where the results seem to be close (cf. Chapter 1, Section 6). Note that the difference here is statistically significant according to a two-tailed Fisher’s exact test (p=0.0377).
Figure 1. First person singular present indicative endings by right context and by region, seventeenth century.

Figure 2. First person singular present indicative endings by right context and by region, eighteenth century.

rises to over 70%. The difference between the proportion of zero forms in vocalic and consonantal right contexts is never more than ten percentage points in these three regions. In Zeeland, however, the pattern is reversed, and the difference between the two
right contexts is also larger. In Zeeland, both vocalic and consonantal right contexts mainly trigger schwa endings, which make up over 60% in the case of a following vowel, and over 70% in that of a following consonant. Focusing on the incoming variant, a difference of 18 percentage points becomes apparent in that zero endings occur in 32% of vocalic right contexts, and in no more than 14% of consonantal right contexts. In consonantal right contexts, the change is thus still incipient, while it is already new and vigorous in vocalic right contexts, according to Nevalainen & Raumolin-Brunberg’s (2003: 55) stages of change (see footnote 3).

Finally, we investigated the interplay of first person endings, right contexts and region in the eighteenth-century main regions subcorpus. Figure 2 confirms the results presented in Tables 4 and 7 above. By the end of the eighteenth century, the written language is levelled to such an extent that both the regional differences and the differences between the two right contexts have largely disappeared. Both contexts trigger about 70% zero forms in Amsterdam, South Holland and Zeeland, while the scores even rise to over 80% in North Holland.

3.3 Left context

Boutkan & Kossmann (1998) describe schwa apocope in historical Dutch as a primarily phonological change and single out one phonological context where schwa is often maintained, viz. following d. Convincing evidence comes from nouns such as schade ‘damage’ and hulde ‘tribute’ referred to above. On the other hand, first person singular present indicative forms of verbs with a stem in t already occurred without final schwa in the Early Middle Dutch period, when schwa apocope was just emerging. For a verb such as laten ‘let’, with the stem lat-, schwa-less forms are attested in the dictionary of Early Middle Dutch covering the thirteenth century (VMNW s.v. laten). This means that we can identify two phonological left contexts, the first slowing down schwa apocope, the latter promoting it.

On the assumption that preceding d is the most conservative context and preceding t the most progressive, we extracted all first person singular present indicative forms of verbs with stems in d and t from the present subcorpora, and categorised the forms according to their endings. Verbs with a d-stem include amongst others bidden ‘pray’, houden ‘hold’, worden ‘become’ and zenden ‘send’. Verbs with t-stems are, for instance, laten ‘let’, zitten ‘sit’, groeten ‘greet’ and wachten ‘wait’. Table 9 presents the results based on the seventeenth-century main regions subcorpus.

The results in Table 9 clearly demonstrate the important effect of left contexts on the distribution of the endings. Stems in t have 92% zero forms, viz. much more than the average seventeenth-century score of 59% (cf. Table 2 above). Stems in d, on the other hand, are very conservative in only attracting 35% zero forms. It should be noted that thepreterito-presens weten ‘know’ is included in the category of stems in t. Excluding weten would, however, not influence the results, despite the fact that weten is fairly frequent. The number of zero forms with stems in t would drop to 442, and the
Table 9. First person singular present indicative endings by left context, seventeenth century.

<table>
<thead>
<tr>
<th>17th century</th>
<th>ø</th>
<th>〈e〉</th>
<th>〈en〉</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>stems in d</td>
<td>40</td>
<td>35</td>
<td>63</td>
<td>55</td>
<td>9</td>
</tr>
<tr>
<td>stems in t</td>
<td>590</td>
<td>92</td>
<td>47</td>
<td>7</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 10. First person singular present indicative endings by left context, eighteenth century.

<table>
<thead>
<tr>
<th>18th century</th>
<th>ø</th>
<th>〈e〉</th>
<th>〈en〉</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>stems in d</td>
<td>55</td>
<td>62</td>
<td>30</td>
<td>34</td>
<td>4</td>
</tr>
<tr>
<td>stems in t</td>
<td>192</td>
<td>91</td>
<td>15</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

total number of forms with stems in t would become 489. The proportion of zero forms would thus be (442/489=) 90%.

Table 10 gives the eighteenth-century results which still show a strong difference. Stems in d remain remarkably conservative, although the proportion of the incoming zero variant has risen to 62%, which is, however, low compared to the 73% zero forms attested overall in the eighteenth century (cf. Table 2).

Turning to the influence of phonological left contexts across region, considerable differences arise. Figure 3 presents the seventeenth-century results, based on Table 9. The overall result of over 90% zero forms with t-stems (Table 9) is confirmed in North Holland, Amsterdam and South Holland. In Zeeland, too, t-stems constitute the most progressive context, attracting the incoming variant in 74% of the tokens. At the same time, schwa-forms remain below 10% in the three most northerly regions, while making up a quarter of the endings following t in Zeeland. With d-stems, the picture is completely different. The overall result of 35% apocope (Table 9) holds for Amsterdam and South Holland, where the incoming variant is therefore outnumbered by the variants with schwa. This is even more the case in Zeeland, where forms with schwa make up (80% 〈e〉 +12% 〈en〉 =) 92%. North Holland, on the other hand, is more progressive than Amsterdam and South Holland and the only region where the incoming variant outnumbers the forms with schwa: d-stems co-occur with zero endings in 65% of the tokens. In sum, the left context is an important variable, exerting the strongest influence in Zeeland. The results also suggest that apocope of final schwa is a regional phenomenon spreading from north to south.
Chapter 9. Apocope of final schwa

Figure 3. First person singular present indicative endings by left context and by region, seventeenth century.

Strong distributional differences such as these may be expected to generate social variation. Before turning to social differences within each region, we discuss the distribution of first person endings across regions in the eighteenth century; see Figure 4.

Figure 4. First person singular present indicative endings by left context and by region, eighteenth century.
Figure 4 shows that apocope of final schwa has progressed both phonologically and regionally. Now Zeeland also scores approximately 90% zero forms with $t$-stems. Stems in $d$ already attracted the incoming variant in about two-thirds of the tokens in seventeenth-century North Holland, a score which is also reached by Amsterdam and South Holland in the eighteenth century. In Zeeland, $d$-stems remain somewhat more conservative, with 50% zero forms and 50% schwa endings. Generally speaking, the eighteenth-century results present a far more levelled picture than the seventeenth-century results.

4. North Holland, including Amsterdam

In this section and the following two, we discuss social variation in the distribution of first person endings, focusing on the four major regions distinguished in the two subcorpora. As we have seen in the previous section, North Holland and Amsterdam generate similar results, although North Holland tends to be slightly more progressive. In this section, we will combine these two regions, which has the advantage of generating many tokens and thus greatly enhancing the reliability of the analyses. This is important for the lower and lower-middle classes in particular, where numbers tend to be low in the individual regions. As South Holland constitutes the transitional zone to Zeeland, which has so far generated very different results, we decided not to include South Holland in the same analyses, despite the fact that it mainly patterned similarly to North Holland in the previous sections.

Table 11 presents the distribution of first person endings in North Holland and Amsterdam in the seventeenth and eighteenth centuries, based on Tables 3 and 4 above. Table 11 shows that the overall distribution of first person endings in North Holland and Amsterdam does not differ greatly in the seventeenth and eighteenth centuries. The proportion of the incoming zero variant is 71% in the seventeenth-century subcorpus and 75% in the eighteenth-century subcorpus, while the proportion of forms with schwa drops from (22+4=) 26% to (17+3=) 20%.

Taking social variation into account and using the seventeenth-century autograph subcorpus, the two periods appear to be quite different after all. Interestingly, apocope of final schwa offers a rare example of the importance of age as a variable. Figure 5 presents the proportion of the incoming zero variant by age and for both periods. It

| Table 11. First person singular present indicative endings, North Holland and Amsterdam diachronically. |
|-------------------------------------------------|------------------------|-----------------|-----------------|-----------------|-----------------|
| North Holland and Amsterdam                      | $\bar{0}$            | $\langle e \rangle$ | $\langle en \rangle$ | Other           | Total           |
| N       | %     | N   | %     | N   | %     | N   | %     | N   | %     |
| 17th century | 1,063 | 71 | 337 | 22 | 59 | 4 | 43 | 3 | 1,502 | 100 |
| 18th century | 835 | 75 | 191 | 17 | 37 | 3 | 55 | 5 | 1,118 | 100 |
should be recalled that we distinguish three age groups, viz. below 30, 30 to 50 years old, and over 50. The black columns representing seventeenth-century usage show that the youngest age group scores just over 70%, and the middle group only slightly less, while the proportion drops to less than 50% in the >50 group. In the eighteenth-century, the picture is entirely levelled, all groups having crossed the 70% line. The seventeenth-century results suggest that apocope of final schwa was a steadily proceeding sound change more strongly affecting younger generations than older generations.

Similar trends appear for the variables of gender and class, that is differences in the seventeenth century and levelling in the eighteenth century. Compare Figure 6 which gives the proportion of the incoming variant by gender and for both periods. In the seventeenth century, women appear to be more progressive than men, using over 70% zero forms, while men score just slightly over 50%. The eighteenth-century columns show more levelling, with both men and women crossing the 70% line, and women even scoring over 80% zero forms. A social class difference is shown for the seventeenth century, represented by the black columns in Figure 7. In the seventeenth century, the lower and middle classes accept the incoming variant in more than 60% of the tokens. In the upper class, however, the proportion of zero endings is just over 50%. Eighteenth-century usage again proves to be much more levelled, all ranks scoring 70% or more.
Figure 6. Proportion of the incoming zero variant by gender, based on the seventeenth-century autograph and the eighteenth-century main regions subcorpora.

Figure 7. Proportion of the incoming zero variant by social class, based on the seventeenth-century autograph and the eighteenth-century main regions subcorpora.
Seventeenth-century usage thus shows both gender and social variation, with men and the upper classes behaving more conservatively than women and the lower and middle classes. In an effort to identify the most conservative writers, we focus on male and female UC writers from this period. Table 12 presents the distribution of the variants in their letters. Numbers for women are fairly low, but the distribution in letters by men is quite telling. The proportion of the incoming zero variant is extremely low at only 43%, while endings with schwa make up no less than (51+2=) 53%. In other words, the new variant is even outnumbered by the schwa variants.

The schwa apocope change in progress was already far advanced in seventeenth-century North Holland and Amsterdam, as is borne out by the overall results. The incoming variant makes up about three-quarters of all tokens. A few social groups could be identified as conservative vis-à-vis apocope of final schwa. The fact that age proved to be an important variable, with >50 writers scoring less than 50% apocope, implies that we are dealing with generational change. This might suggest that the variation in first person endings did not reach the level of social awareness. Research on gender and social variation, however, showed that at least for some writers the variants indexed social values. Upper class men, in particular, were shown to behave conservatively, preferring the historical variants to the incoming zero forms.

5. South Holland

As stated above, we excluded South Holland from the analyses in the previous section because of its status as a transitional zone between Holland and Zeeland. The present section will show that the social effects of variation in first person endings in South Holland are indeed different from those in North Holland and Amsterdam. Table 13 presents the distribution of first person endings in South Holland in the seventeenth and eighteenth centuries, derived from Tables 3 and 4 above.

As in the previous section, the overall results in Table 13 give the impression of a stable situation, with just over 70% incoming zero variants in the seventeenth as well as in the eighteenth century, and 23 to 24% endings with schwa. However, important differences appear between the two periods when social variables are taken into account. Figure 8 shows strong age differences in the distribution of the incoming zero variant.
<table>
<thead>
<tr>
<th>South Holland</th>
<th>o</th>
<th>⟨e⟩</th>
<th>⟨en⟩</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>17th century</td>
<td>308</td>
<td>73</td>
<td>85</td>
<td>20</td>
<td>14</td>
</tr>
<tr>
<td>18th century</td>
<td>291</td>
<td>71</td>
<td>90</td>
<td>22</td>
<td>8</td>
</tr>
</tbody>
</table>

**Figure 8.** Proportion of the incoming zero variant by age, based on the seventeenth-century autograph and the eighteenth-century main regions subcorpora.

Whereas the two younger age groups score fairly highly in the seventeenth century, the >50 group reaches just over 20%. This pattern is similar to the age pattern established in North Holland and Amsterdam (see Figure 5 in Section 4), though the >50 score of only 23% is much lower in South Holland against the 46% score in North Holland and Amsterdam. The eighteenth-century distribution is again much more levelled, as all age groups score well over 60%.

Figure 9 clearly shows that in the seventeenth century the gender differences are less striking in South Holland than in North Holland and Amsterdam. Whereas men from North Holland and Amsterdam scored 52% incoming variants and women 72%, the difference is only 12 percentage points in South Holland, where men score 63% and women 75%. Diachronically, the ongoing change is confirmed by the male scores in Figure 9, where there is a clear rise in the proportion of the incoming variant. The results for letters written by women are less straightforward, in that the proportion actually drops to 60% in the eighteenth century. We will come back to this in Section 7 below. First, we will discuss the social class distribution of the incoming zero variant; see Figure 10.
Figure 9. Proportion of the incoming zero variant by gender, based on the seventeenth-century autograph and the eighteenth-century main regions subcorpora.

Figure 10. Proportion of the incoming zero variant by social class, based on the seventeenth-century autograph and the eighteenth-century main regions subcorpora.
The social differences displayed in Figure 10 appear to be stronger than those established for North Holland and Amsterdam (cf. Figure 7 in Section 4). The LC and LMC have high proportions of the incoming variant in both periods. In the UMC in the seventeenth century, the proportion is much lower, viz. just over 60%, while the UC are most conservative with only 36% zero forms. In the eighteenth century, the distribution is much more levelled, mainly because of a sharp rise in the UC score, but the UC and UMC still seem to lag behind compared to the LC and LMC. Interestingly, the proportion of zero forms even drops a few percentage points in the UMC. As with the reversed gender difference established above, we will come back to this in Section 7 below. An important observation is that in South Holland in the seventeenth century, conservative behaviour with regard to apocope of schwa is mainly associated with the UC and UMC, apart from the generational differences.

As in Section 4, we try to pinpoint the most conservative seventeenth-century writers by focusing on the distribution in the UMC and UC, divided by gender; see Table 14. Numbers are low, so caution is called for. Nevertheless, the results clearly suggest that the most conservative group in North Holland and Amsterdam, viz. UC men (of whom unfortunately no data exist for South Holland), is extended to UC women and UMC men, who score proportions of only 36% and 43%, respectively.

Table 14. First person singular present indicative endings in the UC and UMC, by gender, seventeenth century.

<table>
<thead>
<tr>
<th>South Holland</th>
<th>ø</th>
<th>〈e〉</th>
<th>〈en〉</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>UMC-M, 17th c.</td>
<td>6</td>
<td>43</td>
<td>7</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>UMC-F, 17th c.</td>
<td>10</td>
<td>83</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>UC-M, 17th c.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>UC-F, 17th c.</td>
<td>14</td>
<td>36</td>
<td>24</td>
<td>62</td>
<td>1</td>
</tr>
</tbody>
</table>

6. Zeeland

Finally, we will discuss the results from Zeeland focusing on the social variables. Table 15 gives the overall distribution of the first person variants in Zeeland, derived from Tables 3 and 4 above. As noted before, the seventeenth-century results for Zeeland stand out for the low proportion of the incoming variant, viz. only 23%, while the eighteenth-century results are in line with the results of the other regions. Turning to the social variables, and to age in particular, we again witness strong differences with the other regions; see Figure 11. Unfortunately, the >50 group had to be excluded.
Table 15. First person singular present indicative endings, Zeeland diachronically.

<table>
<thead>
<tr>
<th>Zeeland</th>
<th>ø N</th>
<th>%</th>
<th>〈e〉 N</th>
<th>%</th>
<th>〈en〉 N</th>
<th>%</th>
<th>Other N</th>
<th>%</th>
<th>Total N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>17th century</td>
<td>145</td>
<td>23</td>
<td>431</td>
<td>68</td>
<td>47</td>
<td>7</td>
<td>8</td>
<td>1</td>
<td>631</td>
<td>100</td>
</tr>
<tr>
<td>18th century</td>
<td>143</td>
<td>71</td>
<td>43</td>
<td>21</td>
<td>11</td>
<td>5</td>
<td>5</td>
<td>2</td>
<td>202</td>
<td>100</td>
</tr>
</tbody>
</table>

Figure 11. Proportion of the incoming zero variant by age, based on the seventeenth-century autograph and the eighteenth-century main regions subcorpora.

From further analysis due to lack of data: no results from the >50 group are found in the Zeeland part of the seventeenth-century autograph subcorpus and the eighteenth-century >50 writers generated only 6 first person singular present indicative forms. In the seventeenth century, both remaining age groups have proportions of 20 to 25% incoming variants. At first sight, this may lead to the conclusion that apocope had not yet proceeded to such an extent that age differences can be witnessed. This seems to contrast with the results in North Holland, Amsterdam and South Holland where age proved to be an important factor in the seventeenth century. The importance of age as a factor of change, however, mainly relied on the difference between the two youngest age groups, which scored similarly, and the >50 group (cf. Figure 5 and Figure 8). Thus the missing >50 group in Zeeland prohibits us from drawing a conclusion about the importance of age as a factor in this region in the seventeenth century. In the eighteenth century, the <30 group is somewhat behind the 30–50 group, but the main conclusion is that the change has clearly progressed significantly, both groups reaching approximately 70 to 80% apocope, precisely as in the other regions.
The gender difference attested in seventeenth-century Holland also appears to hold in Zeeland. In the seventeenth-century, represented by the black columns in Figure 12, female letter writers are more progressive than male letter writers. While men show a proportion of just over 20%, women are at 35% zero forms. Figure 12, furthermore, clearly confirms the sharp increase of zero forms from the seventeenth to the eighteenth century. In the eighteenth century, men have progressed to almost 80% apocope, while women have risen but clearly lag behind relative to eighteenth-century men, scoring only 54% zero forms. Remember that we established the same reversed gender difference in eighteenth-century South Holland, which we will discuss further in Section 7. First we will address the social class variable; see Figure 13. Owing to lack of data, we have grouped together the LC and the LMC as well as the UMC and the UC.

Figure 13 indicates the absence of social class variation in the seventeenth-century, which constitutes a remarkable difference from North Holland, Amsterdam and South Holland. In the eighteenth century, the LC and LMC have progressed greatly, almost reaching 80% zero forms. Now variation across social class appears, as the UMC and UC have 65% zero forms. Surprisingly, UMC and UC women behave most conservatively in the eighteenth century; see Table 16.

![Figure 12](image-url). Proportion of the incoming zero variant by gender, based on the seventeenth-century autograph and the eighteenth-century main regions subcorpora.
Figure 13. Proportion of the incoming zero variant by social class, based on the seventeenth-century autograph and the eighteenth-century main regions subcorpora.

Table 16. Endings of the first person singular present indicative in the U(M)C, by gender, eighteenth century.

<table>
<thead>
<tr>
<th>Zeeland</th>
<th>ø</th>
<th>(e)</th>
<th>(en)</th>
<th>Other</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>U(M)C-M, 18th c.</td>
<td>40</td>
<td>70</td>
<td>16</td>
<td>28</td>
<td>1</td>
</tr>
<tr>
<td>U(M)C-F, 18th c.</td>
<td>27</td>
<td>59</td>
<td>11</td>
<td>24</td>
<td>5</td>
</tr>
</tbody>
</table>

In the previous section on South Holland and in this section on Zeeland, we have reported on results that seemed to differ from the results for North Holland and Amsterdam, as well as from the seventeenth-century results for South Holland. We will now offer an overarching interpretation of these results.

7. Interpreting social and regional variation

It should be recalled that in North Holland and Amsterdam, apocope of final schwa had gained a certain level of social awareness, resulting in high proportions of forms with schwa for men and for the UC, and in particular for UC men. In other words, this group of writers behaved most conservatively. In Section 5, we found that conservative
behaviour vis-à-vis apocope of final schwa was extended to UC women and UMC men in South Holland in the seventeenth century. In the eighteenth century, however, women were found to be more conservative than men in South Holland (Figure 9 in Section 5), and the UMC behaved more conservatively than the UC (Figure 10 in Section 5). Similarly, Section 6 showed that eighteenth-century women from Zeeland were more conservative than men (Figure 12), and within the relatively conservative UMC and UC in particular, women were slightly more conservative than men (Table 16). How should we interpret all these observations and conclusions? Why does the same sound change produce such different results in different periods, and in different areas? To find an answer to this, we need to scrutinize the eighteenth-century situation in South Holland and Zeeland, in particular in the UMC and the UC.

Figure 14 presents the proportion of the incoming form across social class and gender, in both periods, combining the data for South Holland and Zeeland. It should be noted that we excluded the seventeenth-century UC men from the analysis due to the scarcity of data (only three tokens) available. Figure 14 shows that UMC men and women as well as UC women score quite conservatively in the eighteenth century, staying within the 50 to 60% range. The eighteenth-century average in both regions was over 70% (cf. Table 13 and 15 above). UC men, on the contrary, are quite progressive, appropriating zero forms in 80% of the tokens. This result contrasts with the conservative behaviour established for seventeenth-century UC men in North Holland and Amsterdam.

![Figure 14](image.png)

**Figure 14.** Proportion of the incoming zero variant by social class and gender, South Holland and Zeeland diachronically, based on the seventeenth-century autograph and the eighteenth-century main regions subcorpora.
The explanation for the varying results is related to the changing sociolinguistic situation. In the seventeenth century, apocope of final schwa progressed in spoken and in written language. We can witness this change in any social group, as all of them allow zero forms into the written code. Some groups of writers, unsurprisingly those with the most writing experience, perceived the historical forms with schwa as the preferred variants, presumably adhering to an age-old writing tradition. This applies to UC men in North Holland and Amsterdam, and in South Holland also to UC women and UMC men. This can be understood because the change had progressed less far in South Holland than in North Holland and Amsterdam, and was consequently raised above the level of social awareness among more letter writers. In the meantime, the LMC and LC produce high proportions of zero forms in North Holland and South Holland, and in the eighteenth century also in Zeeland, bearing witness to the ongoing sound change in the spoken language and thereby also allowing it in their written language. In South Holland and Zeeland in the eighteenth century, the whole UMC as well as UC women remain fairly conservative in their usage, but UC men stand out as progressive. The explanation is that the sociolinguistic landscape had changed dramatically, in that a new norm for the written language had emerged in the early decades of the eighteenth century, a norm which was more in line with the spoken language of the Holland dialect area.

From the sixteenth century onward, first person indicative forms with and without final schwa feature in metalinguistic discourse, bearing witness to the variation prevalent in the spoken language. While some commentators merely mention the existence of two options, others show a preference for either the long form or the form with apocope (Dibbets 1995: 263–266). By the middle of the seventeenth century, it seems that forms with apocope are established as the most common forms, being prescribed in the grammars by Kók (1649) and Leupenius (1653). In the authorized version of the Dutch Bible, the Statenbijbel (1637), however, the historical forms prevail, and at the beginning of the eighteenth century, some commentators, including the influential grammarian Moonen (1706) prescribed forms with schwa in the first person present indicative. Nevertheless, after a long period of variation both in language use and in metalinguistic discourse, schwa-less forms dominate from the early eighteenth century onward, with some grammarians no longer even mentioning the existence of forms with schwa (Simons 2013: 190–193). In sum, whereas metalinguistic discourse had previously described the existing variation, or had even prescribed the historical variants, zero variants were gradually established as the common forms, thus developing into an unambiguous writing norm. Figure 14 suggests that UC men were the first to adopt this change in the sociolinguistic situation, exchanging the old writing practice with schwa for the new norm. It does not come as a surprise that it is the group of writers commonly associated with the highest degree of writing experience that first picks up and conforms to the newly established norm. The result is that we find many zero forms in the LMC and LC, where the written language conforms to the spoken language, and also among UC men, who may of course also follow the spoken
language, but who certainly seem to adhere to the new writing norm, seeing that their most immediate peers, viz. UC women and the UMC in general still largely prefer the historical forms with schwa.

8. Epistolary conventions

Since conventional usage may preserve variants that have become obsolete in more creative language use, we will discuss the interplay of apocope of final schwa and epistolary conventions in the present section. Many of the letters in our corpora contain epistolary formulae, which we have analysed as prefabricated or lexicalised phrases taken as a whole from memory (see Chapters 3, 4 and 5). These might be expected to maintain first person forms with schwa, thereby slowing down the ongoing change. Formulaic language and its effect on apocope of final schwa is the topic of Section 8.1. Another distinguishing feature of the language of letters is the frequent absence of first person subjects which, according to contemporary metalinguistic discourse, might have a conservative effect on the ending of the finite verb.5 We will, therefore, also investigate whether zero subjects have any influence on the distribution of first person endings.

8.1 Formulaic language

Our data comprise a wide variety of verbs, many of which occur only a few times. We, therefore, restricted our investigation to verbs that produced 20 tokens or more. We excluded hebben ‘have’ from the analysis, which is mostly used as an auxiliary, on the assumption that lexical verbs such as hopen ‘hope’ en weten ‘know’ are stronger markers of formulaic sequences. Besides, hebben will be at the heart of Section 8.2. We furthermore distinguish between formulaic and non-formulaic usage: some verbs only occur in formulaic sequences, such as bevelen ‘commend’ (6a), others only occur in non-formulaic parts of the discourse (6b), while still others may be used both in epistolary formulae (6c) and in non-formulaic contexts (6d).

(6) a. jck Beueel ue onder d protec d almoogend godt
   I commend you under the protection the almighty God
   ‘I commend you into the protection of the almighty God’

5. Ten Kate (1723–I: 551; our translation) writes: ‘This last style [i.e. with final schwa] still exists with merchants, when they, for the sake of brevity, leave out the pronoun I in writing, writing zende aen u (mitto tibi), etc’. The phrase zende aen u ‘send you’ may of course be an epistolary formula, but ten Kate clearly suggests a link with the omitted pronoun. The subjectless style may have emerged in business writing, but our data show that it was also widely in use in private correspondence.
b. *dat ick meen dat gij mijn gelt sal stieren*  
that I mean that you me money will send  
‘that I mean that you will send me money’

c. *voors soo wete Ick niet veel te schrijven als*  
further so know I not much to write as  
‘further, I do not know much more to write than’

d. *ick weet dat vl wel nae vijs verlanght als vl eens tus komp*  
I know that you well to fish long as you once home come  
‘I know that you will long for fish by the time you get home’

We will first discuss the seventeenth-century, then the eighteenth-century results. Figure 15 presents the proportion of zero endings used with verbs in non-formulaic contexts, while Figure 16 shows verbs in formulaic contexts.

**Figure 15.** Proportion of the incoming zero variant by (non-formulaic) verb, seventeenth-century main regions

**Figure 16.** Proportion of the incoming zero variant by (formulaic) verb, seventeenth-century main regions
It should be noted that Figures 15 and 16 confirm the importance of phonological left contexts: stems in \( t \) are among the most progressive verbs (\textit{weten} ‘know’, \textit{laten} ‘let’), while stems in \( d \) are among the most conservative ones (\textit{zenden} ‘send’, \textit{bidden} ‘pray’). There is, however, one formulaic verb that is even more conservative than \textit{bidden}, viz. \textit{bevelen} ‘commend’. This could be interpreted as an indication that formulaic contexts as such may, at times, be conservative, and thus maintain schwa endings. Unfortunately, no instances of \textit{bevelen} occur in a non-formulaic context, which excludes the possibility of a comparison of both contexts. Furthermore, it should be noted that most non-formulaic verbs have proportions of 50 to 60% zero forms. Apart from \textit{zenden}, with its \( d \)-stem, \textit{wensen} ‘wish’ and \textit{komen} ‘come’ are also conservative, while \textit{verzoeken} ‘request’ is more progressive. Rather than suggesting differences between formulaic and non-formulaic contexts, these observations suggest verb-specific rates of apocope, or, in other words, lexical diffusion, which relates to phonological conditions, at least in the case of \textit{weten}, \textit{laten} and \textit{zenden}. Most formulaic verbs have slightly smaller proportions of the incoming variant, viz. 40 to 50%. Again, this is an indication that formulaic contexts as such may preserve the historical schwa endings. Since a few verbs occur in both formulaic and non-formulaic contexts, we will focus on these in order to more fully understand the influence of epistolary formulae; see Figure 17.

![Figure 17](image)

\textbf{Figure 17.} Proportion of the incoming zero variant with verbs used in formulaic and non-formulaic contexts, based on the seventeenth-century main regions subcorpus.

Figure 17 shows no difference in the case of \textit{weten} ‘know’: in formulaic as well as non-formulaic contexts, proportions of the incoming variant peak at well over 90%. Similarly, the verb \textit{hopen} ‘hope’, which also occurs with the prefix \textit{ver}-, does not differ greatly in both contexts. The third verb \textit{wensen} ‘wish’ does show a clear difference, but not the expected one: it attracts over 50% zero endings in formulaic contexts, and just
over 30% in non-formulaic contexts, which are, in other words, conservative. Thus the comparison does not prove any conserving effect of formulaic language. What it does suggest is that verbs may behave differently in different contexts.

Summing up, evidence of lexical diffusion appeared in the investigation of a selection of verbs. At the same time, we found some evidence that formulae may behave conservatively vis-à-vis apocope of final schwa, but also evidence of the opposite. We suggest that these phenomena could be captured in the concept of constructional diffusion, discussed in Section 6 of Chapter 8. We will further discuss this with regard to the eighteenth-century data.

Turning to the eighteenth century, Figure 18 presents the proportion of zero endings used with verbs in non-formulaic contexts, and Figure 19 with verbs in formulaic contexts.
contexts. As expected, the proportions are consistently higher than in the seventeenth century. In non-formulaic contexts, all verbs have proportions of 50% or more, most verbs even cross 60% or 70%, while *weten* ‘know’ and *laten* ‘let’ cross 90%. Three non-formulaic verbs remain below 60%, viz. *zenden* ‘send’, *komen* ‘come’ and *verlangen* ‘long, desire’. In formulaic contexts, *verlangen* is fairly conservative as well, along with *wensen* ‘wish’ and *(ver)blijven* ‘stay, remain’. Is this conservative behaviour linked to the verbs as such and do we see only lexical diffusion or do the verbs behave differently in different contexts and can we establish constructional diffusion? To examine these issues, we compare the proportions of verbs appearing in formulaic as well as non-formulaic contexts; see Figure 20.

![Figure 20. Proportion of the incoming zero variant with verbs used in formulaic and non-formulaic contexts, based on the eighteenth-century total subcorpus.](image)

As can be seen in Figure 20, small differences exist between the formulaic and non-formulaic uses of *weten* and *laten*, though the more important observation is that these verbs rise above 90% in both contexts. The verbs *verzoeken* ‘request’ and *verlangen* ‘long’ have identical proportions in both contexts, but the scores differ greatly, *verzoeken* reaching no less than 80% apocope while *verlangen* has only slightly over 50%. This difference once more indicates that lexical diffusion is at work. The scores for *hopen* ‘hope’, *wensen* ‘wish’ and above all *(ver)blijven* ‘stay, remain’ differ in line with our expectations, viz. that formulaic sequences co-occur with the schwa variants to a greater extent than non-formulaic sequences. The verb *(ver)blijven*, in particular, shows the most obvious difference between both contexts. Of the non-formulaic use of *(ver)blijven* in the eighteenth-century subcorpus, 11 first person tokens occur, 8 of which (73%) are realised without an ending. The verb appears much more frequently in formulaic contexts (72 tokens), where it is predominantly used with a schwa ending (39
tokens (54%) with final e and 5 tokens (7%) with final en), as opposed to only 26 (36%) zero ending forms. The formulaic contexts are all instances of the closing formula exemplified in (7a–c), which exerts a very conservative influence on the verb form.

(7) a. *vaart wel god zegent uEd, ik blijf met agting en tedere*  
fare well God blesses you I remain with respect and tender  
dankleving UEDW Dienaar & vriend  
gratitude your servant and friend  
‘fare well, God bless you. With respect and tender gratitude, I remain your servant and friend’

b. *myn lief heer mede moet ik hfsbreken en ik blyve ve en*  
my dear here with must I break off and I remain your  
liefdragend vrouwe  
loving wife  
‘my dear, with this I have to finish, and I remain your loving wife’

c. *de groeten nis aan alle goede vrienden en bekenden in haast*  
the greeting to all good friends and acquaintances in haste  
*Ik verblyfe uld getrouwe zoon*  
I remain your faithful son  
‘greetings to all good friends and acquaintances. In haste. I remain your faithful son’

Interestingly, this closing formula is much more frequent in the eighteenth century (72 tokens) than in the seventeenth century (10 tokens). This means that it not only provides a conservative context in which schwa endings can survive; with its increasing usage in the eighteenth century, the formula even promotes the historical schwa forms.

8.2 Zero subjects

As noted above, a distinguishing feature of the language of letters is the frequent absence of first person subjects. A possibly preserving effect of zero subjects on the ending of the finite verb is somewhat implicitly suggested by contemporary metalinguistic discourse (see footnote 5). Examples of zero subject context are the seventeenth-century (8a) and eighteenth-century (8b), the former with a schwa-less finite verb, the latter with a schwa form.

(8) a. *gelijck vdt vl brijef vorstan heb dat vl schrijef dat*  
as from your letter understood have that you write that  
‘as I have understood from your letter that you write that ...’

b. *ter wiel hier daagelijks zoo veel scheepe koomen van amsterdam*  
while here daily so many ships come from Amsterdam
Letters as Loot

_Hebbe den_ 12 _deezes alhier een brief ontfangen_

have the 12 this+gen here a letter received

‘because so many ships from Amsterdam come here every day. I have received a letter here on the 12th of this month’

We investigated the effect of the presence of the subject by focusing on one verb, viz. the highly frequent verb _hebben_ ‘have’, thus keeping the data as homogenous as possible. We extracted all schwa-less forms (_heb_) and forms with schwa (_hebbe, hebben_) from the subcorpora, judged the person and number (_hebbe and hebben_ can be plural forms of all persons), and then grouped them into two categories, viz. overt subject and zero subject. Tables 17 and 18 present the results.

**Table 17.** Endings of the first person singular present indicative of _hebben_ ‘have’ with and without a subject, based on the seventeenth-century main regions subcorpus.

<table>
<thead>
<tr>
<th></th>
<th>ø</th>
<th>⟨e⟩</th>
<th>⟨en⟩</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>overt subject</td>
<td>357</td>
<td>44</td>
<td>385</td>
<td>48</td>
</tr>
<tr>
<td>zero subject</td>
<td>44</td>
<td>25</td>
<td>129</td>
<td>72</td>
</tr>
</tbody>
</table>

Table 17 shows that, in the seventeenth century, _hebben_ attracts 44% zero endings when used with an overt first person subject, and (48+8=) 56% schwa forms. When the subject is omitted, however, the proportion of zero endings drops to 25%, while the proportion of schwa forms rises to 75%. We may conclude that zero subjects favour schwa forms. This also applies to the eighteenth century (Table 18), where overt subjects trigger 83% zero endings, while zero subjects only attract 48% zero endings. The epistolary convention of omitting first person subjects thus appears to slow down schwa apocope, providing a conservative context where schwa endings are maintained.

Since _hebben_ generates such a large number of tokens in both periods, we decided to combine the results for _hebben_ with the main external variables to find out whether the difference between overt and zero subjects would hold. Figure 21 presents the proportion of zero endings in the first person singular of _hebben_ with overt and zero
subjects, split by region and time. In North Holland, Amsterdam and South Holland in both periods, and in Zeeland in the eighteenth century, zero endings outnumber schwa endings with overt subjects (the black columns). When the subject is omitted, however, the proportion of zero endings consistently drops, in the seventeenth century to approximately 30%, in the eighteenth century to 40 to approximately 60%. The one exception is Zeeland in the seventeenth century, where the difference is actually inverted, and where the important observation is that zero endings hardly occur, neither with overt subjects nor with zero subjects.

The difference between overt subjects and the schwa preserving context of zero subjects also holds in every age group in both periods: the proportion of zero endings is consistently lower when the subject is omitted, as can be seen in Figure 22. The smallest difference is found in the seventeenth-century age group of >50, where schwa apocope hardly occurs at all.

Two things stand out in Figure 23, which presents the results across gender and time. Whereas eighteenth-century men and seventeenth-century women present the now expected pattern of fewer zero endings when the subject is omitted, seventeenth-century men and eighteenth-century women do not differ greatly in their proportions of zero endings depending on the presence of the subject. Seventeenth-century men simply exhibit little schwa apocope, while eighteenth-century women clearly prefer zero endings in both contexts.

Finally, Figure 24 combines the results with social class and time. For expository reasons, we have grouped the LC and LMC together, as well as the UMC and the UC. Figure 24 shows that the difference between overt subjects and zero subjects also holds across social class and time. Zero subjects generate fewer zero endings and preserve
Figure 22. Zero endings in the first person singular present indicative of hebben ‘have’ with and without a subject, diachronically by age, based on the seventeenth-century autograph and the eighteenth-century total subcorpora.

Figure 23. Zero endings in the first person singular present indicative of hebben ‘have’ with and without a subject, diachronically by gender, based on the seventeenth-century autograph and the eighteenth-century total subcorpora.

schwa forms. The difference is less striking in the UMC and UC in the seventeenth century than elsewhere. Remember that the UMC and UC were quite conservative vis-à-vis the ongoing change, especially UMC and UC men in Holland.
Figure 24. Zero endings in the first person singular present indicative of *hebben* ‘have’ with and without a subject, diachronically by social class, based on the seventeenth-century autograph and the eighteenth-century total subcorpora.

A clear difference emerges between overt subjects and zero subjects in both periods, zero subjects preserving the schwa forms. Although we are primarily interested in identifying factors influencing the choice of first person endings, an intriguing question is why zero subjects would preserve schwa endings. After all, person and number are uniquely indicated by first person singular forms. The most common finite verb endings in our letters are e/ø – t – t in the singular (e.g. lopeloop – loopt – loopt ‘walk – walk – walks’), and en – t – en in the plural (e.g. lopen – loopt – lopen ‘walk’). This implies that both the schwa and the zero ending are unique to the first person singular. One could even argue that zero endings are ‘more unique’, as plural forms are often written without final -n and are thus identical to singular schwa forms. Morphology, therefore, cannot account for the schwa-maintaining effect of zero subjects. The only way to explain this, in our view, is to assume an abstract construction [zero first person subject + verb-e] that specifies the epistolary convention of the finite verb receiving schwa when the subject is omitted. Thus again we would find constructional diffusion. Why this conventionalised pattern would exist, is a question beyond the scope of this study. What is interesting from our sociolinguistic perspective, is that the difference between overt and zero subjects holds across region, gender, age and social class. This means that, broadly speaking, letter writers of all ages and social backgrounds, both men and women, from all regions represented in our corpus, were aware of this

---

6. Obviously, all this applies to the written language. In spoken language, plural forms were probably even much more frequently pronounced without final -n, while especially in the singular, many different paradigms must be assumed, with zero endings in second and third person forms, and t-endings in the first person, which we have also seen in the present chapter.
convention, once more bearing witness to the strong supraregional patterns found in these private letters.

9. Conclusions

In this chapter, we discussed apocope of final schwa in first person finite verb forms of the present indicative. In the first part of the study, we confirmed suggestions in the literature that apocope of final schwa was a phonological change spreading from north to south. We have identified conservative and progressive left contexts, while arguing that the influence of the right context was less important. Apocope of final schwa may have been promoted by hiatus avoidance in spoken language; in written language this effect is only modest. In the eighteenth century, most of the variation disappeared, giving way to much more uniformity.

Apocope spread from north to south, broadly speaking from Holland to Zeeland, but we have found evidence allowing for a more detailed picture of spreading from North Holland to Amsterdam, South Holland, and then to Zeeland. Interestingly, for this change age has proved to be an important variable in most regions. Whereas we have not found many generational differences in our studies so far, we have been able to show clear generational patterns in the case of schwa apocope, attesting to the fact that the variation and change remained unnoticed by many language users. Nevertheless, it rose above the level of social awareness for other groups, which appears from the social class and gender variation that we have also established. The social differences followed the regional diffusion of the change, in that they first surfaced in Holland, and only later in Zeeland.

In the seventeenth century, phonological factors determined the results for Zeeland, where the change was just taking off. To the north, in South Holland, Amsterdam and North Holland, the change had progressed, and social variation could be established. The change had probably advanced most strongly in North Holland and Amsterdam, where the language community had by then largely switched to schwa-less forms, while the conservative group of UC men still clung to schwa endings. The change was still in flux in South Holland, where awareness of variation was more widespread, and conservative writers were also found among UC women and UMC men.

In the eighteenth century, the overall results for North Holland and Amsterdam were similar to those of the previous period, but the change had largely vanished from social awareness, as no social variation could be established. Schwa forms were still in use, however, for instance in epistolary conventions. Social variation did still exist in South Holland and Zeeland, where we found remarkable patterns, in that women were more conservative than men, and the UMC writers more than the UC, mainly because UC men adopted schwa-less forms to a considerable degree. The difference from seventeenth-century North Holland and Amsterdam, where UC men were the most
conservative group, was explained with reference to the changing sociolinguistic situation. In the seventeenth century, an opposition existed between spoken language with progressive apocope and written language, where old writing practices showed either variation or a clear preference for schwa forms. In the eighteenth century, a new and schwa-less writing norm had developed, and UC men were among the first to adopt this new convention.

In the final section of this chapter, we looked into the influence of epistolary conventions on apocope of final schwa. We were able to show that different verbs exhibit very different proportions of zero endings, suggesting that lexical diffusion is an important factor. Interestingly, we also found evidence that a particular verb may behave differently depending on the context in which it is used. It may be conservative in formulaic contexts and moderately progressive in non-formulaic contexts, such as *(ver)* blijven in the eighteenth century, or conversely. This is proof that the context in which a verb occurs co-determines its morphology and that, in other words, constructional diffusion plays a role. Another letter-writing convention is the zero subject, and we demonstrated that zero subjects provide a schwa-maintaining context. This result holds across all external variables, indicating that this particular convention was well-known to most letter writers in the western Low Countries.
Chapter 10

Clausal and local negation

1. Introduction

In the present chapter, we will focus on a topic that is among the best-researched areas in the history of Dutch and in historical linguistics in general, viz. negation. Specifically, we will focus on the final stage of the so-called Jespersen's cycle from bipartite (1a) to single negation (1b) in seventeenth-century Dutch, taking into account both clausal and local negation. The bipartite negation in (1a) consists of the preverbal negator *en*, which dates back to Old Dutch, and the newer postverbal negator *niet*. Note that the terms preverbal and postverbal refer to the position vis-à-vis the finite verb taken by the negators in main clauses. Examples (1c–e) contain other, less frequent negators such as, in present-day spelling, *geen* ‘no’ (1c), *niemand* ‘nobody’ (1d) and *nooit* ‘never’ (1e).

(1) a. ick *en* verget v *niet* in mijn gebedt
    I NEG forget you NEG in my prayer
    ‘I won’t forget you in my prayer’

b. ick vergeet het sterven mijn leven *niet*
    I forget the dying my life NEG
    ‘I won’t forget the dying in all my life’

c. ick *en* sal mijn moeder *gen* droefheijdt aendoen
    I NEG shall my mother NEG sadness cause
    ‘I will not hurt my mother’

d. *Iegenwoordigh jsser* niemant die maelt
    at present is-there NEG who cares
    ‘at present nobody cares’

e. *tabago* dat sij nu *noijt* meer sien sal
    Tobago that she now NEG anymore see will
    ‘Tobago, which she will not see anymore now’

In the literature on the changes in negation patterns in Dutch (i.a. van der Horst & van der Wal 1979; de Haan & Weerman 1984; Burridge 1993; Hoeksema 1997; Goss 2002; Rutten et al. 2012; Vosters & Vandenbussche 2012), a wide variety of internal and external variables have been discussed. After having introduced the subcorpora and the overall results in Section 2, we will concentrate on what appear to be the most important
variables from the perspective of the research tradition referred to, viz. region and construction type. It has been argued by many (i.a. van der Horst & van der Wal 1979; Burridge 1993; Rutten et al. 2012; Vosters & Vandenbussche 2012) that the change from bipartite to single negation spread from the north to the south, affecting different semantico-syntactic construction types at a different pace. In Sections 3 and 4, we will focus on the regional and constructional patterns found in our subcorpora. It has also been suggested (i.a. Hoeksema 1997; Goss 2002; Vosters 2011) that specific verbs and/or the phonetic context may have promoted the negation change. This will be investigated in Sections 5 and 6. In Section 7, we will discuss to what extent different types of complexity may have been an important factor, focusing on the length of the utterance and the number of verbs in the verbal cluster. Section 8 discusses the overall results of the social variables rank, gender and age. In Section 9, we will combine the most important internal variable (construction type) with the social variables, while concentrating on two regions (Amsterdam and Zeeland).

Throughout the chapter, we will discuss only seventeenth-century results. Elsewhere, using a selection of 100 private letters from the eighteenth century, we reported that we had found only 12 instances of bipartite negation and over 600 tokens of single negation in this subset (Rutten et al. 2012). While it is in itself interesting to note that bipartite negation still occurred in private letters from the late eighteenth century, and also in the Northern Netherlands, we have excluded the eighteenth-century material from further analysis due to the extremely low frequency of bipartite negation overall.

2. The subcorpora and the overall results

As has been said, in the present chapter we will focus primarily on the seventeenth century. There are strong indications that the changes in the expression of clausal and local negation spread from the north to the south of the language area, and were conditioned by internal factors. We will, therefore, first discuss regional and internal factors, for which we will use the entire collection of private letters from the seventeenth century, viz. 549 letters written by 424 different writers. This collection includes established autographs, established non-autographs and letters of which the autograph status is uncertain (cf. Chapter 1, Section 6). We call this subcorpus the private subcorpus, as it comprises all private letters from the seventeenth century available for research at the time of writing the present monograph. The regional subcorpus used in Chapter 9, comprising 454 letters from the four major regions, North Holland, Amsterdam, South Holland and Zeeland, is part of this private subcorpus. The private subcorpus used here, however, also contains letters from other regions such as Friesland and Flanders. For the research on social variables, we confined ourselves to the 260 established autographs written by 202 different writers. 219 of these letters could be linked to a social class index (cf. Table 2 in Chapter 1, Section 6, partly repeated here as Table 1).
In the following sections, we will discuss a range of independent variables influencing the choice of negation type. In order to be able to do so, we extracted all negations from the subcorpora used, mainly by searching for post-verbal negators in various spellings, including niet ‘not’, geen ‘no’, niemand ‘nobody’, nimmer ‘never’ and nooit ‘never’. The overall results based on the private subcorpus are presented in Table 2.

As can be seen in Table 2, niet and geen are by far the most frequent post-verbal negators, making up 2,251 tokens of the 2,307 negations. The distribution of negation types is approximately two-thirds single and one-third bipartite, which scores differ slightly, however, in the case of the less frequent negators niemand, nimmer and nooit. Finally, it should be noted that the corpus search also resulted in 28 double negations, which were excluded from the analyses in the following sections.1

3. Regional patterns

A possible regional patterning of single and bipartite negation in the history of Dutch has been suggested in the literature (e.g. van der Horst & van der Wal 1979; Paardekooper

---

1. The very low number of the supposedly colloquial double negation may come as a surprise. It is unclear whether this low number reflects the frequency of double negation in the spoken language or is the result of a strong proscription against double negation of the type ‘two negatives affirm’, which repeatedly occurs in the Dutch normative tradition as well as in German and English (cf. Nevalainen & Rutten 2012).
Burridge (1993) compared texts from Holland in the north of the language area to texts from Brabant, which is located in the mid-south of the language area. Her corpus ranged from 1300 to 1650, and for the sixteenth and seventeenth centuries mainly comprised published and edited texts, written by well-educated men (Burridge 1993: 266–268). She demonstrated that bipartite negation lasted longer in the south than in the north, as she found relatively more bipartite negation in texts from Brabant than in texts from Holland. In fact, she found 100% single negation in main clauses in texts from Holland from around 1650, and no less than 99% single negation in subordinate clauses (Burridge 1993: 192). We will come back to differences between various semantico-syntactic patterns such as main clauses and subordinate clauses in Section 4.

In the eighteenth and nineteenth centuries, the overall impression is that bipartite negation may still have been used in the spoken language of the north, while it is rare in the written language. Nevertheless, even grammarians and literary authors from the north used it sometimes (van der Horst 2008: 1573). Vosters & Vandenbussche (2012) show that contrary to the near absence of bipartite negation in eighteenth- and nineteenth-century northern Dutch, the variant was still fairly well established in the written language of southern regions such as Flanders, Brabant and Limburg, even in administrative and legal sources from the 1820s.

The north-south differences observed for past periods are confirmed by present-day dialect data from the syntactic atlas of the Dutch dialects (SAND). Maps 48b, 49a, 49b and 50a in the SAND show that bipartite negation in main clauses is maintained in only Flemish dialects (i.e. French-Flemish, West-Flemish, East-Flemish) in present-day Belgium and the north of France, while map 50b shows that bipartite negation in subordinate clauses is maintained in a larger area, covering not just the Flemish dialect areas but also the Brabant area in Belgium, with moreover a handful of attestations in Belgian Limburg. Nowadays, to sum up, bipartite negation appears almost exclusively in the southern and south-western parts of the language area.

Turning to the regional distribution in our seventeenth-century corpus, we expect that in the event of salient regional differences, these will be in line with the foregoing, expecting, for instance, more bipartite negation in letters from Zeeland than in letters linked to Holland. Considering the fact that Burridge (1993) established an almost exclusive use of single negation in texts from Holland from around 1650, we do not expect to find many bipartite negations in letters linked to Holland, especially since the letters in our corpora are more recent, dating from the 1660s and 1670s.

From the 549 letters of the private subcorpus, which were assigned to regions (as explained in Chapter 1, Section 5), we extracted all negations and allocated these to the appropriate regions. Table 3 shows the proportion of single and bipartite negation in each region. In the most northern major region, North Holland, single negation is by far the dominant variant (88%). In the city of Amsterdam, located in the south of North Holland, the incoming variant is used in two-thirds of all negations. To the south of Amsterdam, in the provinces of South Holland and Zeeland, single and bipartite negation
Table 3. Single and bipartite negation by region.

<table>
<thead>
<tr>
<th>Region</th>
<th>Single N</th>
<th>Single %</th>
<th>Bipartite N</th>
<th>Bipartite %</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>North Holland</td>
<td>385</td>
<td>88</td>
<td>53</td>
<td>12</td>
<td>438</td>
</tr>
<tr>
<td>Amsterdam</td>
<td>523</td>
<td>67</td>
<td>253</td>
<td>33</td>
<td>776</td>
</tr>
<tr>
<td>South Holland</td>
<td>134</td>
<td>49</td>
<td>142</td>
<td>51</td>
<td>276</td>
</tr>
<tr>
<td>Zeeland</td>
<td>285</td>
<td>52</td>
<td>262</td>
<td>48</td>
<td>547</td>
</tr>
<tr>
<td>Flanders</td>
<td>15</td>
<td>42</td>
<td>21</td>
<td>58</td>
<td>36</td>
</tr>
<tr>
<td>Other</td>
<td>41</td>
<td>91</td>
<td>4</td>
<td>9</td>
<td>45</td>
</tr>
<tr>
<td>Unknown</td>
<td>118</td>
<td>62</td>
<td>71</td>
<td>38</td>
<td>189</td>
</tr>
<tr>
<td>Total</td>
<td>1,501</td>
<td>65</td>
<td>806</td>
<td>35</td>
<td>2,307</td>
</tr>
</tbody>
</table>

are more or less equally frequent. The number of letters and thus of negations from the (West) Flemish region is limited compared to the previous regions, but it is interesting to note that the proportion of single negation now drops to 42%, bipartite negation being the dominant variant here. The small category Other includes the eastern region of Gelderland (3 tokens), but also, importantly, Friesland (7 tokens) and letter writers originating from German- and Norwegian-speaking areas (35 tokens). In Frisian, High and Low German and Norwegian, the shift from bipartite to single negation is dated earlier than in Dutch (cf. Elspaß & Langer 2012; Nevalainen & Rutten 2012). On the assumption that morphosyntactic imposition of L1-structures such as single negation may occur in the Dutch of these letter writers, the early shift to single negation in these languages is clearly reflected in the high proportion of single negations in the category Other, viz. 91%. The letters labelled Unknown could not be linked to any specific region since no relevant information on the writers was found, as in the case of letters of which the writers are unknown. The proportions for single and bipartite negation in the letters of unknown regional origin tie in neatly with the overall pattern, viz. about two-thirds single and one-third bipartite negation.

The results in Table 3 show that the change from bipartite negation to single negation was, perhaps first and foremost, a regional phenomenon. From North Holland in the north-west of the language area to West Flanders in the south-west, the proportion of single negation steadily drops, and the proportion of bipartite negation concomitantly rises. The results, therefore, confirm the suggestions made in previous studies, and also link up with the status of bipartite negation in present-day dialects. Focusing on seventeenth-century Holland, however, the results are somewhat surprising, as the letters from Holland, and in particular from Amsterdam and South Holland contain many more bipartite negations than expected on the basis of Burridge’s investigation of Hollandic Dutch from around 1650. Whereas single negation prevailed or was even the only pattern in her data, bipartite negation occurs in 33% of the negations in letters linked to Amsterdam, while taking up no less than 51% in letters linked to South...
Holland. There is, in other words, a considerable difference between the published and edited texts from around 1650 used by Burridge, and the more recent yet more conservative private letters used for the present study.

4. Constructional patterns

The type of negation also depends on the type of clause, and in particular a correlation has been established between negation and word order. Earlier studies argued that the choice of negation is strongly determined by the semantico-syntactic environment or the construction type. Focusing on the position of the finite verb, Burridge (1993) in her study of texts from Holland and Brabant from the period 1300–1650 distinguished between three syntactic environments: V1, V2 and V-final. In V1-sentences, the finite verb is sentence-initial as in imperatives and yes/no questions (2a). V2 is the common word order of Dutch main clauses (2b), while V-final refers to the word order in subordinate clauses (2c). V1, V2 and V-final are the three basic word order possibilities in Modern Dutch, and these three options have been at the core of research on negation in Dutch. The following examples, taken from our corpus, all contain bipartite negation; the two elements are in boldface.

\[(2)\]  
\[\text{a. } ende en verkert altijt in geen herbergen}\]  
\[\text{and } \text{NEG be always in no taverns ‘and never go to taverns’}\]

\[\text{b. maer godt en heeft het soo niet gewilt}\]  
\[\text{but God NEG has it so not wanted ‘but God did not want it (to be) this way’}\]

\[\text{c. dat het de koninck niet hebben en wilt}\]  
\[\text{that it the king not have NEG want ‘that the king does not want it’}\]

Van der Horst & van der Wal (1979) distinguished two more environments, viz. subject-verb inversion, triggered by the fronting of another part of the discourse (2d), and local negation (2e). To these five environments, we added constructions where niet and geen are not adverbial negators to other parts of the discourse, but definite pronouns with the meaning ‘nothing’ and ‘nobody’ respectively, and as such act as a part of the discourse itself, viz. subject in (2f). Van der Horst & van der Wal (1979) found solely single negations in this context, and therefore did not discuss this environment. However, we did encounter variation of negation type in this context.

\[(2)\]  
\[\text{d. soo en konde ick ul niet naerder schrijven}\]  
\[\text{so NEG could I you not more write ‘Thus I could not write you more’}\]
e. waer op ick tegenwoordich noch niet meer als 6000 op betaelt
where on I to this day yet not more than 6,000 on paid
en hebben
NEG have
‘of which to this day I have paid not more than 6,000’

f. alhier en passert niet van merito
here NEG passes nothing of merit
‘here nothing happens worth mentioning’

A solid result of van der Horst & van der Wal (1979) and Burridge (1993), repeatedly confirmed in later studies, including the recent research by Vosters & Vandenbussche (2012), was that V1 clauses are the first to adopt single negation, while subordinate clauses with V-final constitute the most conservative construction type, at least in the seventeenth century.²

To assess the validity of the claims about the importance of the construction type, we allocated all 2,307 negations to their respective environments, distinguishing between main clauses and subordinate clauses, main clauses with V1, main clauses with subject-verb inversion, local negation and constituent negation where niet and geen are discourse parts. Note that sentences with niemand ‘nobody’ are structurally similar to these, and that the constituent category therefore comprises negations both with niemand and with niet and geen in their pronominal discourse functions. Table 4 presents the results.

It is immediately clear from Table 4 that the construction types identified in earlier studies exert a considerable influence on the choice of negation type. There are semantico-syntactic constraints on the distribution of the variants, which suggests that

<table>
<thead>
<tr>
<th>Construction Type</th>
<th>Single</th>
<th>Bipartite</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>V1</td>
<td>118</td>
<td>89</td>
<td>15</td>
</tr>
<tr>
<td>Local</td>
<td>157</td>
<td>82</td>
<td>35</td>
</tr>
<tr>
<td>Constituent</td>
<td>110</td>
<td>77</td>
<td>32</td>
</tr>
<tr>
<td>Main clause</td>
<td>501</td>
<td>67</td>
<td>249</td>
</tr>
<tr>
<td>Inversion</td>
<td>157</td>
<td>56</td>
<td>122</td>
</tr>
<tr>
<td>Subclause</td>
<td>458</td>
<td>56</td>
<td>353</td>
</tr>
<tr>
<td>Total</td>
<td>1,501</td>
<td>65</td>
<td>806</td>
</tr>
</tbody>
</table>

² In earlier periods, e.g. in Middle Dutch, subordinate clauses constituted a progressive clause type, displaying a fairly high proportion of single negation (van der Horst & van der Wal 1979: 8–10).
the incoming variant spreads from one construction type to another (constructional diffusion, cf. Chapter 8, Section 6). In particular, the correlation between the three main word orders and negation is evident. Whereas V1-constructions are very progressive, displaying 89% single negation, ‘normal’ V2 main clauses have 67% single negation, while V-final subordinate clauses prove to be conservative with only 56% single negation.

Two features that are less progressive than V1-clauses and more progressive than main clauses are local negation and constituent negation. Note that 35/142 examples of constituent negation contain the element *niemand*, which co-occurred with single negation in 81% of the instances (Table 2 in Section 2). This appears to be in line with the 77% single negation established for constituent negation in Table 4. In other words, *niet* and *geen*, when used as pronouns with a function such as subject or object, pattern in accordance with *niemand* and thus behave more progressively than when used as adverbal negators (65% single negation, see Table 2). Furthermore, main clauses with subject-verb inversion are quite conservative, scoring only 56% single negation. Burridge (1993: 207–208) argued that main clauses with inversion need not be considered separately from other main clauses, since the word order is similar when it comes to the finite verb, i.e. V2. Van der Horst & van der Wal (1979), however, claimed that in the Middle Dutch period main clauses with inversion were among the first constructions to switch from bipartite to single negation. Similarly, Vosters & Vandenbussche (2012: 356), focusing on nineteenth-century southern Dutch, found that main clauses with inversion behaved more progressively than normal main clauses and subclauses. Against this background, the results in Table 4 may come as a surprise. We would have expected main clauses with inversion to either pattern as other main clauses do or behave even more progressively than normal main clauses, but instead they behave as conservatively as subclauses.

To more fully understand the surprising results for inverted clauses, we will focus on the three main construction types involved, viz. main clauses, main clauses with inversion and subclauses, and examine whether the results for inversion hold across region. In Table 5, we present the results, divided per region, limiting ourselves to the four main regions offering a considerable number of data.

**Table 5. Single negation by construction type and region.**

<table>
<thead>
<tr>
<th></th>
<th>Main clause</th>
<th>Inversion</th>
<th>Subclause</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>North Holland</td>
<td>120</td>
<td>87</td>
<td>36</td>
</tr>
<tr>
<td>Amsterdam</td>
<td>174</td>
<td>68</td>
<td>58</td>
</tr>
<tr>
<td>South Holland</td>
<td>45</td>
<td>47</td>
<td>13</td>
</tr>
<tr>
<td>Zeeland</td>
<td>100</td>
<td>57</td>
<td>34</td>
</tr>
</tbody>
</table>
Table 5 gives the proportion of single negation for each construction type. First of all, we note that the pattern of decreasing proportions of single negation when traveling from the north to the south generally holds for each construction type, which is clear evidence that both region and construction type are major factors influencing the choice of negation. Furthermore, inverted clauses appear to be slightly more conservative than the other two types in North Holland, but we would rather stress that the proportion of single negation is 80% or more in all construction types. This means that the change is reaching completion in North Holland, and we, therefore, do not wish to attribute too much importance to the differences between main clauses, inverted clauses and subclauses. In Amsterdam, subclauses are more conservative than main clauses, while inverted clauses pattern in almost perfect accordance with normal main clauses. This suggests that main clauses with inversion can be treated as other main clauses, both representing V2 word order. In South Holland and Zeeland, however, we do find differences between inversion and the other types. In South Holland, inverted main clauses are much more conservative than both normal main clauses and subclauses. In Zeeland, inverted clauses are not as conservative as subclauses, but are still substantially more conservative than main clauses. In other words, the fact that main clauses with inversion prove to be surprisingly conservative in Table 4 may be largely due to their relatively conservative behaviour in South Holland and Zeeland. How do we have to interpret this?

What distinguishes inverted clauses from normal main clauses is the fronting of a part of the discourse other than the subject, which moves to the position immediately following the finite verb, viz. the third position. We may hypothesise that particular fronted elements tend to maintain bipartite negation. Different elements can be fronted among which we will distinguish three basic options. The first is the fronting of a referential part of the discourse that primarily functions at the level of the clause such as the direct object *brandewijn* ‘brandy’ in (3a). The second option is similar, but the part of the discourse that is fronted is a clause in itself, e.g. a temporal subclause (3b). The third option concerns structural elements of grammaticalised discourse of which *voorst* ‘further’ as in (3c) is among the most frequent ones (cf. Chapter 3, Section 6.2). These elements do not have a referential function at the clause level, but a connective function at the level of textual coherence.

(3)  

a. *brandewijn* *can* ick *niet* *verkoopen*  
   *brandy* can I not sell  
   ‘I cannot sell brandy’  

b. *toen* ick *desen briei* *schreef* en *hebbe* ick *noch den domenij* *niet*  
   when I this letter wrote neg have I yet the minister not  
   *ghesijen ghehad*  
   seen had  
   ‘when I wrote this letter, I had not seen the minister yet’
c. voorst en wet ick u myn lyeff man nyet meer te schryv
‘further, my dear husband, I do not know anything more to write’

We coded all the main clauses with inversion in accordance with these three basic options, focusing on the deviating regions South Holland (41 inverted clauses) and Zeeland (71 inverted clauses). Table 6 presents the results.

Although the absolute numbers are fairly low, we can formulate some conclusions, which can be used as hypotheses in future research. When a part of the discourse is fronted that primarily functions at the level of the clause, the proportions of single negation are 50% in South Holland and 69% in Zeeland, resembling or even surpassing the proportions of single negation in normal main clauses, which were 47% and 57%, respectively. In other words, if in a normal main clause one of the discourse parts is fronted to the position of the subject, this does not influence the choice of negation or has a progressive influence on it. The conservative behaviour established above for main clauses with inversion is due to the two other types of inversion, involving subclauses and connectives. In these cases, the proportions of single negation drop below the level of main clauses, reaching extremely low values of about 30% single negation in many cases. What these two types of inversion have in common is that the fronted elements function at the level of the larger discourse rather than at the level of single main clause. Fronted subclauses expand the inverted main clause, which is turned into a complex sentence. Connectives create coherence between the main clause with inversion and the preceding discourse. In sum, there appears to be a correlation between the maintenance of bipartite negation and discourse elements above the level of the simple clause.

The link between negation and discourse coherence is not a new phenomenon. It has been observed previously that negators may grammaticalise into coherence markers or indicators of specific syntactic constructions. In the Middle Dutch period, bipartite negation is the most common pattern, but the single preverbal negation *ne/en* lives on in specific constructions where it often signals conditionality, as the clitic in example (4), attached to the subject *ic ‘T’.*

(4) *ic salre varen, in blive doet* (13th c.; from van der Horst 2008: 518)
I will-there go i-neg stay dead
‘I will go there, unless I die’

<table>
<thead>
<tr>
<th></th>
<th>South Holland</th>
<th></th>
<th>Zeeland</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>clause-level part</td>
<td>5</td>
<td>50</td>
<td>10</td>
<td>69</td>
</tr>
<tr>
<td>subclause</td>
<td>3</td>
<td>30</td>
<td>11</td>
<td>56</td>
</tr>
<tr>
<td>connective</td>
<td>5</td>
<td>24</td>
<td>13</td>
<td>35</td>
</tr>
</tbody>
</table>

Table 6. Single negation in main clauses with inversion by region and fronted element.
In (4), the historically negative element functions as a subordinate conjunction with the meaning ‘if not, unless’. In similar cases, Burridge (1993: 181–182) speaks of paratactic negation, with the negator fulfilling the function of clause linkage. Similar though much more recent are twentieth-century examples from the dialect of Katwijk aan Zee in the province of South Holland. This dialect has been reported to maintain bipartite negation, though only in subclauses. Bipartite negation is, however, on the decline, and in this process *en* has acquired the possibility of appearing in positive subclauses as well, where it merely functions as an indicator of subordination; see (5).

\[
\text{(5) ik loof dat ik negen jaar en was} \quad (20\text{th c.; cf. de Vink 2004: 116})
\]

\[
\begin{align*}
\text{I believe that I nine years neg was} \\
\text{‘I believe that I was nine years old’}
\end{align*}
\]

The results in Table 6 suggest that negative elements also took up the function of discourse marker or indicator of specific syntactic constructions in the seventeenth century, at the height of the change from bipartite to single negation in written northern Dutch.

5. **Lexical-semantic patterns**

Lexical-semantic influences on clause negation have also been suggested. Both the type of negator and the type of negated verb have been related to specific types of negation. In the present section, we will first investigate possible differences between the negators *niet* and *geen* (5.1), and then discuss the influence of the type of verb (5.2).

5.1 **The negators *niet* and *geen***

Vosters (2011: 427) hypothesises a difference between *niet* and *geen*. In the change from bipartite to single negation, *niet* and *geen* have become the negators proper to language users. The highly frequent form *niet* may have been the forerunner in this development, representing the default or neutral means of clause negation and, as a result, *niet* would be expected to behave progressively. Vosters’ corpus of nineteenth-century Dutch, however, did not yield any significant differences between *niet* and *geen* in the selection of negation types. Another view, founded on the idea of cognitive entrenchment as a result of high frequency (e.g. Bybee 2006, 2010), would predict that *niet* behaves conservatively. The pattern *en ... niet*, which is more common than *en ... geen*, would be maintained as a result of the preserving effect associated with highly frequent patterns. The two opposing hypotheses following from this, viz. that *niet* is either progressive or conservative when compared to *geen*, will be investigated by comparing *niet* and *geen* in the three basic word orders, viz. V1, V2 main clauses and V-final subclauses, as well as in the total score for all construction types (cf. Section 4). The results are in Table 7.
Table 7. Single and bipartite negation with niet and geen.

<table>
<thead>
<tr>
<th></th>
<th>Single</th>
<th></th>
<th>Bipartite</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>V1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>niet</td>
<td>103</td>
<td>90</td>
<td>11</td>
<td>10</td>
<td>114</td>
</tr>
<tr>
<td>geen</td>
<td>14</td>
<td>82</td>
<td>3</td>
<td>18</td>
<td>17</td>
</tr>
<tr>
<td>Main clauses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>niet</td>
<td>394</td>
<td>67</td>
<td>191</td>
<td>33</td>
<td>585</td>
</tr>
<tr>
<td>geen</td>
<td>104</td>
<td>66</td>
<td>54</td>
<td>34</td>
<td>158</td>
</tr>
<tr>
<td>Sub clauses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>niet</td>
<td>334</td>
<td>55</td>
<td>274</td>
<td>45</td>
<td>608</td>
</tr>
<tr>
<td>geen</td>
<td>114</td>
<td>60</td>
<td>75</td>
<td>40</td>
<td>189</td>
</tr>
<tr>
<td>All construction types</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>niet</td>
<td>1,187</td>
<td>65</td>
<td>643</td>
<td>35</td>
<td>1,830</td>
</tr>
<tr>
<td>geen</td>
<td>275</td>
<td>65</td>
<td>146</td>
<td>35</td>
<td>421</td>
</tr>
</tbody>
</table>

In V1-clauses, both niet and geen show over 80% single negation. Although niet appears to be slightly more progressive at 90% single negation, we would rather say that the change is nearing completion with both niet and geen. The results, however, considering the fairly low number of tokens in the case of geen, do not provide an empirical base from which firm conclusions can be drawn. This is different with main clauses, subclauses and in the total scores for all construction types, where the number of tokens is considerably higher. Contrary to the two opposing hypotheses, no significant differences between niet and geen can be established.³ It cannot be excluded that a more fine-grained analysis of clauses with niet and geen may reveal differences in the choice of negation type. On the basis of the results in Table 7, however, it does not seem likely that the type of negator exerts a significant influence on the choice of negation type, be it progressive or conservative.

5.2 Type of verb

Burridge (1993: 180) argues that specific verbs maintained single preverbal negation longer than others in the change from single preverbal to bipartite negation in the earlier Middle Dutch period. The observation mainly applies to ‘common usage verbs’ such as modals and say, do, know and speak. Burridge also suggests that their high frequency promoted their conservative behaviour. Hoeksema (1997: 143–145) continues this line of research, arguing that it is rather negative polarity-sensitive verbs that

³. Differences between niet and geen within the construction types as well as across all construction types are not statistically significant according to a series of Fisher’s exact tests.
maintain the single preverbal negator *en* or *ne*. He relates this finding to Tieken-Boon van Ostade’s (1987) result that it was verbs most likely to appear in negative or interrogative contexts that retained postverbal negation in eighteenth-century English, when most other verbs preferred the new system with *do*-support. In the history of English, it has also been claimed that high frequency verbs display conservative behaviour vis-à-vis negation patterns (Bybee & Thompson 2000).

Extending these results to the change from bipartite to single negation in Dutch, we could hypothesise that negative polarity-sensitive verbs were conservative, displaying relatively high proportions of bipartite negation. Vosters (2011: 477) indeed found some evidence of negative polarity-sensitive verbs behaving conservatively. In the present section, we investigate whether conservative behaviour related to the type of verb can be established in our corpus of seventeenth-century Dutch.

One problem then is to decide what counts as a negative polarity-sensitive verb. The first and most obvious candidate is *hoeven*, which indeed is a negative polarity item as are its counterparts in English (*need*) and German (*brauchen*) (cf. van der Wouden 1994: 10; Falkenberg 2001: 79). Furthermore, Hoeksema (1994, 1997) mentions the psychological verbs *care* and *matter*, but these do not have frequent counterparts in our corpus. Vosters extends the category of psychological verbs to include *weten* ‘know’, *kennen* ‘know’ and *zeggen* ‘say’, and in his data, these are among the most conservative verbs. So are *mogen* ‘may’ and *zullen* ‘shall’, which suggests that modals are also relatively conservative (cf. Hoeksema 1997). Of the German negative polarity verbs listed by Falkenberg (2001), Dutch counterparts often do not appear in our corpus; only *hesitate* and *dare* are found in our letters.

All in all, we have three hypotheses, which are not entirely compatible. Firstly, high frequency verbs are conservative. Secondly, auxiliaries and especially modals are conservative. Thirdly, negative polarity verbs, including psychological verbs, are conservative (*hoeven* ‘need’, *durven* ‘dare’, *weten* ‘know’, *kennen* ‘know’, *zeggen* ‘say’). The incompatibility lies in the fact that modals and negative polarity verbs need not be frequent.

We investigated the occurrence of single and bipartite negation with each single verb in the corpus, limiting ourselves to verbs that provided at least 6 tokens so as to enhance the reliability of the results. Figure 1 presents the results. The individual verbs with their frequencies in the corpus are located on the y-axis. The x-axis gives the proportion of bipartite negation occurring with each verb.

It should be recalled that the average proportion of bipartite negation is 35% (cf. Table 2). While a large number of verbs are to be found with approximately 30 to 40% bipartite negation, many other verbs diverge from this pattern. Some reach proportions of bipartite negation of over 50 or even over 60% (*komen* ‘come’, *verlaten* ‘leave’, *geloven* ‘believe’, *krijgen* ‘get, receive’, *doen* ‘do’), others drop to about 20% or even lower (*hopen* ‘hope’, *maken* ‘make’, *laten* ‘let’, *gaan* ‘go’). The divergent patterns suggest that there are lexical constraints on the use of single and bipartite negation. However, these are not related to verb type frequency. Frequent verbs such as *zullen* ‘shall’, *zijn* ‘be’ and *kunnen* ‘can’ all score about one third bipartite negation, and whereas *hebben* ‘have’ may appear to be somewhat progressive (29% bipartite negation), *weten* ‘know’ is quite
Figure 1. Bipartite negation by verb type.

conservative (45% bipartite). This enumeration also shows that auxiliaries and/or modals are not necessarily conservative, to which we may add that while *mogen* ‘may’ is progressive (23% bipartite), *doen* ‘do’ is conservative (68% bipartite). Finally, the results do not induce us to think that negative polarity verbs and psychological verbs
prefer bipartite negation. The negative polarity verbs *hoeven* ‘need’ (31% bipartite), *durven* ‘dare’ (31% bipartite) and *twijfelen* ‘hesitate’ (40% bipartite) have normal values of bipartite negation. The psychological verb *weten* ‘know’ diverges from the average and is indeed conservative (45% bipartite). It is unclear what counts as a psychological verb, but if we extend the category to include, for instance, progressive verbs such as *hopen* ‘hope’ (14% bipartite) and *vergeten* ‘forget’ (22% bipartite), and conservative verbs such as *denken* ‘think’ (50% bipartite) and *geloven* ‘believe’ (57% bipartite), we again conclude that these divergent results do not confirm the hypothesis. Even if we limit the category to the verbs semantically closest to *weten*, i.e. *denken* en *geloven*, we cannot conclude that it is negative polarity sensitivity that determines conservative behaviour, as there are other verbs at least as conservative as these three, viz. *komen* ‘come’ (51% bipartite), *verlaten* ‘leave’ (57% bipartite), *krijgen* ‘get, receive’ (62% bipartite) and *doen* ‘do’ (68% bipartite).

In sum, whereas the results suggest lexical-semantic patterns in the choice of negation type, these do not seem to be related to verb type frequency nor to negative polarity or auxiliaries and modals. We admit that detailed analyses of particular items may still reveal lexical-semantic patterns, while also pointing out that this is the first time that verb types and negation types have been compared on such a large number of data.

Intrigued by the lexical differences, we decided to examine more closely the extreme verb scores, that is progressive verbs where the proportion of bipartite negation drops below 25% (*hopen* ‘hope’, *maken* ‘make’, *laten* ‘let’, *gaan* ‘go’, *vergeten* ‘forget’, *mogen* ‘may’), and conservative verbs where the proportion peaks at over 45% (*doen* ‘do’, *krijgen* ‘get, receive’, *geloven* ‘believe’, *verlaten* ‘leave’, *komen* ‘come’, *believen* ‘please’, *denken* ‘think’, *gelieven* ‘please’, *staan* ‘stand’). Against the background of the results in Section 4, we examined the constructions in which these verbs appear, focusing on the three basic word order possibilities, viz. V1, V2 main clauses and subclauses (see Table 8).

The results in Table 8 cast even more doubt on the importance of lexical-semantic patterns. As the Table shows, the conservative behaviour of the upper group of verbs may be due to the fact that they are mostly used (61%) in the conservative construction type of the subclause. The progressive verbs, on the other hand, are remarkably often used in V1-constructions (39%). It is interesting, of course, that some verbs are commonly used in subclauses, and others in V1 clauses. A verb such as *krijgen* ‘get, receive’ appears 25 times in a subclause, many of which are complaints *dat ick geen briefen uan u en krigh* ‘that I do not get any letters from you’. A verb such as *laten* ‘let’ strongly prefers V1-clauses, most of which are hortatives. The negation patterns, however, appear to be linked to the construction types, not to the verb types, so that the extreme verb scores in Figure 1 are probably largely due to the constructions in which the verbs appear, conforming earlier research as well as our results as reported in Section 4.
Table 8. Distribution of conservative and progressive verbs by construction types.

<table>
<thead>
<tr>
<th></th>
<th>V1</th>
<th>Main clause</th>
<th>Subclause</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conservative verbs</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>doen</em></td>
<td>3</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td><em>krijgen</em></td>
<td>0</td>
<td>8</td>
<td>25</td>
</tr>
<tr>
<td><em>geloven</em></td>
<td>0</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td><em>verlaten</em></td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td><em>komen</em></td>
<td>2</td>
<td>13</td>
<td>33</td>
</tr>
<tr>
<td><em>believen</em></td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td><em>denken</em></td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td><em>gelieven</em></td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td><em>staan</em></td>
<td>0</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Total N</td>
<td>11</td>
<td>44</td>
<td>86</td>
</tr>
<tr>
<td>Total %</td>
<td>8</td>
<td>31</td>
<td>61</td>
</tr>
</tbody>
</table>

| **Progressive verbs** |    |             |           |
| *hopen*            | 0  | 5           | 0         |
| *maken*            | 3  | 1           | 2         |
| *laten*            | 27 | 5           | 9         |
| *gaan*             | 2  | 6           | 7         |
| *vergeten*         | 11 | 2           | 4         |
| *mogen*            | 0  | 8           | 17        |
| Total N            | 43 | 27          | 39        |
| Total %            | 39 | 25          | 36        |

6. Phonological patterns

It has been suggested in the literature that phonological restrictions are involved in the use of bipartite negation. Burridge (1993: 194–196) distinguishes three phonetic environments that promote deletion of the negator *en/ne* and thus prefer the single negation pattern: subordinate clauses with an infinitive preceding the finite verb (6a), main clauses beginning with the adverb *dan* ‘then’ or the indefinite pronoun *men* ‘one’ (6b), and main clauses with a plural subject ending in -en (6c). The following Middle Dutch examples, into which we inserted ø to indicate the expected position of the second negator, are taken from Burridge (1993: 195–196, examples 24, 27, 28).

(6) a. *dat helpt wel den ghenen die siin spise niet*  
*that helps well the one who his meal NEG*  
*verduwen ø mach*  
digest NEG can  
*(Holland, c. 1300)*  
‘that greatly helps those who cannot digest their meals’
b. \textit{men ø sal den saffraen niet tevele besighen} (Holland, c. 1300)  
\(\text{one NEG shall the saffron NEG too-much use} \)  
‘don’t use too much saffron’

c. \textit{die boonen ø sien niet goed te verduwen} (Holland, c. 1350)  
\(\text{the beans NEG are NEG good to digest} \)  
‘beans are hard to digest’

The argument appears to be that the left context may haplogologically promote deletion of the historical negator. The left context is always \(-n\), preceded by what may be a schwa (6a, c), [ɛ] (6b) or [α] as in \textit{dan}. Hoeksema (1997: 141–143) also refers to haplogology and establishes a progressive effect of preceding \textit{men} in both Middle Dutch and Early Modern Dutch, but not of preceding infinitives in subordinate clauses. Goss (2002: 176–178), following de Haan & Weerman (1984), states that the loss of \textit{en} is triggered by phonologically similar left contexts in seventeenth-century texts from Holland. Vosters (2011: 471–473), who investigated nineteenth-century Dutch, established a modest effect of preceding \(-n\) and a stronger effect of preceding \(-en\). In his data, mostly \textit{men} and infinite verbs in \(-en\) appear to promote deletion, not nouns in \(-en\) as in example (6c).

We investigated the left context of the (expected position of the) historical negator in all 2,307 tokens of negation in our corpus, distinguishing five different left contexts: vowels as in \textit{wij ‘we’} in (7a), consonants (but not \(\langle n \rangle\)) as in \textit{tijt ‘time’} in (7b), \(\langle n \rangle\) (but not \(\langle en \rangle\)) as in \textit{man ‘man’} in (7c), \(\langle en \rangle\) (but not the lexeme \textit{men}) as in (7d), and finally \textit{men} as in (7e).

\begin{itemize}
  \item \textit{want wij en zijn het niet weerdijgh vanwegen onse menijgh}  
  \(\text{for we NEG are it NEG worthy for our mani} \)  
  \textit{vuldijghie sonden}  
  \(\text{fold sins} \)  
  ‘for we are not worthy of this for our many sins’
  \item \textit{want Jck geen tijt en hadden}  
  \(\text{for I NEG time NEG had} \)  
  ‘for I had no time’
  \item \textit{die man en wilde het gelt niet geeve}  
  \(\text{that man NEG wanted the money NEG give} \)  
  ‘that man did not want to give the money’
  \item \textit{capmessen met houte heften en zijn niet goet}  
  \(\text{chopping-knives with wooden hafts NEG are NEG good} \)  
  ‘chopping-knives with wooden hafts are not good’
  \item \textit{men en weet niet ofte wij het lant sullen mogen houden}  
  \(\text{one NEG knows NEG whether we the land shall may keep} \)  
  ‘one does not know whether we will be allowed to keep the land’
\end{itemize}
Table 9. Single and bipartite negation by left context.

<table>
<thead>
<tr>
<th></th>
<th>Single</th>
<th></th>
<th>Bipartite</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>V</td>
<td>242</td>
<td>61</td>
<td>155</td>
<td>39</td>
<td>397</td>
</tr>
<tr>
<td>C</td>
<td>807</td>
<td>60</td>
<td>528</td>
<td>40</td>
<td>1,335</td>
</tr>
<tr>
<td>〈n〉</td>
<td>55</td>
<td>62</td>
<td>34</td>
<td>38</td>
<td>89</td>
</tr>
<tr>
<td>〈en〉</td>
<td>166</td>
<td>66</td>
<td>84</td>
<td>34</td>
<td>250</td>
</tr>
<tr>
<td>men 'one'</td>
<td>23</td>
<td>96</td>
<td>1</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>1,293</td>
<td>62</td>
<td>802</td>
<td>38</td>
<td>2,095</td>
</tr>
</tbody>
</table>

Excluding both V1-clauses, where the left context is a clause boundary, and ambiguous examples from the analysis, we were able to assign 2,095 negation tokens to the five different left contexts. Table 9 presents the results.

The results in Table 9 do not provide any evidence that phonology as such has an important effect on the distribution of single and bipartite negation. The proportion of single and bipartite negation is roughly 60 and 40% respectively, be the preceding context a vowel or a consonant. Importantly, the same patterns holds when the preceding consonant is 〈n〉, and even when the preceding word ends in 〈en〉. The results are, therefore, not in line with the results obtained by Burridge and Vosters, while they do match the observations by Hoeksema. The strongest left context predictor of single negation is clearly the lexeme men, which co-occurs only once with bipartite negation (see example (7e)), and 23 times with single negation. The corpora investigated by Burridge and Vosters comprise older and younger texts, while Hoeksema claims that his database spans the entire Middle Dutch and Early Modern Dutch periods.4 Obviously, different phonological effects can take place at different times and places. The robust numbers in Table 9 and the size of our corpus, however, do not lead us to assume any other influence than the haplological effect of men.

7. Complexity effects

With the growing literature on linguistic complexity in mind (e.g. Kusters 2003; Dahl 2004; Trudgill 2011), we may wonder whether complexity constrained the use of bipartite negation, which in itself is more complex than single negation on the assumption that complexity depends on the number of words or morphemes needed to express a specific meaning. Goss (2002), for instance, has argued that the change from bipartite to single negation in Holland was at least in part the result of a process of simplification related to immigration and the concomitant dialect contact, leading

speakers to opt for the more simple and less marked type of negation. Seeing that negation is tightly connected to the verb and word order (Section 4), the question also arises of whether the complexity of the verbal cluster correlates with the negation type, specifically whether more complex verbal clusters promote deletion of one of the negators, and thus promote single negation. In addition, there is evidence that distance correlates with complexity (e.g. Gibson 2000), which leads us to hypothesise that a greater distance between the two negators may also promote deletion of one negator.

In the present section, we investigate the possible influence of complexity on the choice of negation, mainly in order to establish whether this can be developed into a fruitful approach in historical sociolinguistics. We, therefore, limit ourselves to a restricted data set, comprising only subordinate clauses with niet as the main negator (608 tokens). This data set was chosen for two reasons. Firstly, in subclauses the verbs tend to cluster together towards the end of the clause, while in main clauses the finite verb is in second position and the other verb forms appear clause-finally. This means that verbal cluster complexity is easier to measure, and in a less ambiguous way, in subordinate clauses than in main clauses. Second, niet ‘not’ and geen ‘no’ behave very differently syntactically, since geen prototypically occurs in attributive positions in front of nouns. Although no differences between niet and geen were established in the selection of negation types (Section 5), their different syntactic behaviour influences the length of utterances. We therefore focused on the negation niet. In what follows, we will first look into the influence of distance on the choice of negation (Section 7.1), and then turn to verbal cluster complexity (Section 7.2).

7.1 Proximity

To assess the influence of distance on the choice of negation, we used what we call the proximity value. The proximity value refers to the distance between the first negator (i.e. niet) and the finite verb, the two elements that are present in any of the 608 tokens. Therefore, the proximity value is at least 2, as in example (8a), where the two relevant items, viz. the negator and the finite verb are in boldface.

(8) a. omdat wij het niet wiste
   because we it NEG knew ‘because we did not know it’

In the bipartite counterpart in (8b), the proximity value is 3.

(8) b. alsoo gij niet en weet
   because you NEG NEG know ‘because you do not know’

With more material intervening in between the first negator niet and the finite verb, the proximity value easily increases to, for instance, 5 in (8c).
Excluding a few ambiguous examples, we were able to compute the proximity value of 597 subordinate clauses with niet. In Table 10, we compare the proximity value $n$ for single negation with the proximity value $n+1$ for bipartite negation, as clauses with bipartite negation are consistently one token longer than clauses with single negation (cf. examples (8a) and (8b)).

Table 10 shows an increase in the proportion of single negation when the proximity value rises, from a 50–50 distribution of single and bipartite negation in the case of fairly short subordinate clauses to approximately 70 to 80% single negation in longer clauses. We interpret this as indicating that longer and in that sense more complex sentences prefer deletion of the en negator, thus promoting single negation. One reason to be somewhat cautious about these data, however, is that the token count decreases to eventually only a handful of examples with proximity values of 6, 7 or higher, which may obviously skew the results. Furthermore, the results for the proximity values 2, 3 and in the case of bipartite negation 4 are surprising, in that the proportion of single negation is only about 50%, which is even slightly lower than the overall distribution of 56% single and 44% bipartite negation established for subordinate clauses in general (see Table 4 in Section 4). This suggests that short clauses are relatively conservative, and against the background of the much higher frequency of relatively short clauses in comparison to longer clauses, a better description of the constraint at hand may in fact be that short clauses tend to be conservative, rather than longer clauses tending to be progressive. In the case of short clauses, however, the proximity value mainly counts the negator(s) and the verb(s), in the absence of intervening material (cf. (8a) and (8b)). But if there is only the verbal cluster apart from the negator(s), the question rises whether it could be the complexity of the verbal cluster itself that determines the negation type. This will be addressed in Section 7.2.

Table 10. Single and bipartite negation by proximity values.

<table>
<thead>
<tr>
<th>Proximity value</th>
<th>Single negation</th>
<th>Proximity value</th>
<th>Bipartite negation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>2</td>
<td>126</td>
<td>50</td>
<td>3</td>
<td>50</td>
</tr>
<tr>
<td>3</td>
<td>108</td>
<td>51</td>
<td>4</td>
<td>49</td>
</tr>
<tr>
<td>4</td>
<td>54</td>
<td>70</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>64</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>83</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>&gt; 6</td>
<td>6</td>
<td>75</td>
<td>&gt; 7</td>
<td>25</td>
</tr>
<tr>
<td>Total</td>
<td>324</td>
<td>54</td>
<td>273</td>
<td>46</td>
</tr>
</tbody>
</table>
7.2 Verbal cluster complexity

The complexity of the verbal cluster refers to the number of verb forms present. In the 597 subordinate clauses with niet (cf. Section 7.1), the verbal complexity values 1, 2, 3 and 4 occur, as in the examples (9a-d). The verb forms are in boldface.

\[(9)\]

\[a. \textit{also het niet kan} \]
because it not can
‘because it cannot happen’

\[b. \textit{alsoo ick tot mijn voorgestelde begeerte niet komen can} \]
because I to my proposed desire not come can
‘because I will not be able to come [to Holland] as [would be] in conformity with my desire’

\[c. \textit{dat de schipper vl niet wilde laten gaen} \]
that the skipper you not wanted let go
‘that the skipper did not want to let you go’

\[d. \textit{dat het niet geberght sou konne warde} \]
that it not stored would can become
‘that it would not be possible to store it’

The large majority of the 597 tokens have a complexity value of 1 (263 tokens) or 2 (300 tokens). The complexity value 3 was allocated to 31 clauses, and value 4 to only 3 clauses. There are strong differences in the choice of negation between subordinate clauses that appear to be linked to the complexity value. Figure 2 plots the complexity values (CV1 etc.) on the x-axis and the proportion of subordinate clauses with single and bipartite negation on the y-axis.

![Figure 2. Single and bipartite negation by complexity value.](image-url)
There appears to be a strong correlation, in that single negation increases and bipartite negation decreases when the complexity of the verbal cluster rises. This strongly suggests that more complex verbal clusters prefer deletion of the *en* negator, thus promoting single negation. Interestingly, the overall proportion of approximately two-thirds single negation corresponds to complexity value 2. When the complexity value is 1, the pattern is almost reversed, no less than 58% of the subordinate clauses selecting bipartite negation. This means that clauses with only one verb form, i.e. the finite verb, tend to be conservative, and maintain bipartite negation. The pattern appears to be more pronounced than the pattern established in Section 7.1, viz. that short clauses tend to be conservative. Since short clauses and clauses with only one verb form are very similar categories, we decided to cross-tabulate complexity values 1 and 2 with the most frequent proximity values, thereby comparing, as before, proximity value *n* for single negation with the proximity value *n+1* for bipartite negation; see Table 11.

With complexity value 2, the proportion of bipartite negation drops below 50% irrespective of the proximity values. The upper part of Table 11, however, shows that there is one particular context that clearly resists *en*-deletion to the extent that bipartite negation occurs in no less than 63% of the tokens, viz. when only a simplex verb is used (CV1) that is immediately preceded by the negator(s). When the length of the utterance, i.e. the proximity value, increases, or when the complexity of the verbal cluster increases, the proportion of single negation rises. This is evidence that complexity as such affects the choice of negation, and that this may be more important than the kind of complexity. Single negation is used more often in longer clauses than in shorter clauses, and more often in clauses with complex verbs than in clauses with simplex verbs, pushing the historical bipartite pattern back to short clauses with simplex verbs.

Table 11. Cross-tabulation of complexity values and proximity values.

<table>
<thead>
<tr>
<th>Complexity value 1</th>
<th>Complexity value 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proximity value</td>
<td>%</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>2 single</td>
<td>37</td>
</tr>
<tr>
<td>3 single</td>
<td>46</td>
</tr>
<tr>
<td>4 single</td>
<td>50</td>
</tr>
<tr>
<td>5 single</td>
<td>48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Proximity value</th>
<th>%</th>
<th>Proximity value</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 single</td>
<td>62</td>
<td>3 bipartite</td>
<td>38</td>
</tr>
<tr>
<td>3 single</td>
<td>53</td>
<td>4 bipartite</td>
<td>47</td>
</tr>
<tr>
<td>4 single</td>
<td>83</td>
<td>5 bipartite</td>
<td>17</td>
</tr>
<tr>
<td>5 single</td>
<td>62</td>
<td>6 bipartite</td>
<td>38</td>
</tr>
</tbody>
</table>
8. Social patterns

There is abundant evidence that, in the seventeenth century, negation rose highly on the scale of social awareness in the Netherlands. Holland-based literary authors such as Hooft (1581–1647) and Vondel (1587–1679), who had shown variation of bipartite and single negation in earlier texts, switched to an almost exclusive use of single negation from about 1640 onward (van der Wal 1990: 64; van der Wouden 1998). As we can date this change almost to the year, it must have been a conscious change, that is, from above the level of social awareness. It should be noted that contemporary literary figures such as Cats (1577–1660) and De Brune (1588–1658), who both originated from the province of Zeeland, maintained both variants in their writings throughout their lifetime (van der Wal 1990: 63–64). From both the seventeenth and the eighteenth centuries, there are also metalinguistic texts bearing witness to the fact that negation had become a sociolinguistic marker or even a stereotype. Only a few years after Hooft’s and Vondel’s switch to single negation, the grammarian Leupenius (1653) explicitly rejected bipartite negation as illogical. Later linguistic commentators likewise condemned bipartite negation. In 1686, for instance, Vollenhove characterised it as solemn or formal (van der Horst 2008: 1299). And well into the eighteenth century, the Holland-based schoolmaster van Belle (1748) argued that two negations result in an affirmation, a cliché that can also be found in the German and English metalinguistic traditions (Nevalainen & Rutten 2012). From van Belle’s comments, we may infer that around 1750 bipartite negation was still in use, although we hardly find any instances in contemporary printed texts. What is more, bipartite negation only rarely occurs in written Dutch from the eighteenth century onward (van der Horst 2008: 1573, 1941), and if it occurs it is (much) less frequent than single negation (cf. Vosters & Vandenbussche 2012), while it survived in many dialects up to the present day. This virtual disappearance from the written code indicates the markedness of bipartite negation.

Against this background, which is, however, largely founded on literary and metalinguistic sources, it seems likely that negation is subject to social variation. Using the autograph subcorpus, we investigated possible correlations between negation types and social variables, starting with social class variation. Table 12 presents the distribution of single and bipartite negation across social rank.

<table>
<thead>
<tr>
<th></th>
<th>Single</th>
<th>Bipartite</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>LC</td>
<td>33</td>
<td>59</td>
<td>23</td>
</tr>
<tr>
<td>LMC</td>
<td>155</td>
<td>60</td>
<td>102</td>
</tr>
<tr>
<td>UMC</td>
<td>427</td>
<td>67</td>
<td>212</td>
</tr>
<tr>
<td>UC</td>
<td>93</td>
<td>70</td>
<td>40</td>
</tr>
<tr>
<td>Total</td>
<td>708</td>
<td>65</td>
<td>377</td>
</tr>
</tbody>
</table>
The two lower ranks produce approximately 60% single negation, while the proportion rises to 70% in the upper class. The differences are small, but the gradual rise of the incoming variant, perfectly in line with what could be expected on the basis of the markedness of the feature, suggests that negation was indeed subject to social variation. For similar reasons, gender effects may be expected. As we have argued in previous chapters, familiarity with written language or so-called writing experience varied according to social rank as well as gender. It is more than likely that those who were more familiar with written language, were also among the most progressive language users in the case of negation. Table 13 presents the results across gender. The difference shown in Table 13 is again small, yet in the expected direction, in that male writers more frequently adopt the incoming single negation variant.

Cross-tabulation of the data for rank and gender gives the results as presented in Figure 3.

Table 13. Single and bipartite negation by gender.

<table>
<thead>
<tr>
<th></th>
<th>Single</th>
<th></th>
<th>Bipartite</th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>507</td>
<td>66</td>
<td>263</td>
<td>34</td>
<td>770</td>
</tr>
<tr>
<td>F</td>
<td>284</td>
<td>58</td>
<td>207</td>
<td>42</td>
<td>491</td>
</tr>
<tr>
<td>Total</td>
<td>791</td>
<td>63</td>
<td>470</td>
<td>37</td>
<td>1,261</td>
</tr>
</tbody>
</table>

Figure 3. Single (and bipartite) negation by social class and gender.

5. The results are not statistically significant according to a series of Fisher’s exact tests.
6. The result is statistically significant according to a Fisher’s exact test (p=0.005).
The social range from about 60% in the lower classes to 70% in the upper class (Table 12) is expanded in both directions, both up and down, in the case of male letter writers. Lower-class men score 50% single negation, which rises to no less than 96% among upper-class men. The score for UC men, in particular, (71 tokens) is highly interesting, as it clearly indicates the social values attached to the different negation patterns. In the case of women, the pattern is less clear-cut. Female writers in the three lower ranks do not seem to differ very much from each other, nor from their male peers, with scores of just over and just below 60%. In the UC, surprisingly, the score for women (62 tokens) drops to 40% single negation, in stark contrast with the male pattern as well as with the overall social pattern established above. On closer inspection, it appears that no less than 30 out of 37 tokens of bipartite negation are produced by female writers allocated to South Holland. This overrepresentation of women from South Holland may obviously have skewed the results. Finally, the combination of the two variables as in Figure 3 suggests that it is particularly among UC men that bipartite negation has become a marked feature.

Our results of the age variable, as presented in Table 14, also suggest that negation had risen above the level of social awareness. The change from bipartite to single negation was clearly not progressing gradually, affecting younger generations to a greater extent than older generations. Instead, we find that all three age groups are affected to the same extent. The differences are very small, and are in line with the overall proportion of about two-thirds single negation.7

In this section, we have studied social variation in the entire autograph subcorpus, abstracting away from, for instance, regional and constructional differences. Yet we have established strong regional and constructional patterns in Sections 3 and 4. It, therefore, seems likely that the social evaluation of variants differs from region to region, and perhaps also from construction to construction type. In the next section, we will again address the issue of social variation, taking into account especially regional and constructional differences.

### Table 14. Single and bipartite negation by age.

<table>
<thead>
<tr>
<th>Age Range</th>
<th>Single N</th>
<th>Single %</th>
<th>Bipartite N</th>
<th>Bipartite %</th>
<th>Total N</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–30</td>
<td>314</td>
<td>60</td>
<td>212</td>
<td>40</td>
<td>526</td>
</tr>
<tr>
<td>30–50</td>
<td>375</td>
<td>65</td>
<td>200</td>
<td>35</td>
<td>575</td>
</tr>
<tr>
<td>50+</td>
<td>63</td>
<td>62</td>
<td>39</td>
<td>38</td>
<td>102</td>
</tr>
<tr>
<td>Total</td>
<td>752</td>
<td>63</td>
<td>451</td>
<td>37</td>
<td>1,203</td>
</tr>
</tbody>
</table>

7. The results are not statistically significant according to a series of Fisher’s exact tests.
9. Social patterns in Amsterdam and Zeeland

In the present section, we will concentrate on two regions that provide a large amount of data and that show considerable differences according to Table 3 in Section 3, viz. Amsterdam and Zeeland. Within these two regions, we will compare V2 main clauses to V-final subordinate clauses, while excluding clauses with *men* ‘one’ in front of the expected position of the historical negator *en*. In this way, we are able to monitor the three main variables of region, construction type and *men* (see Sections 3, 4 and 6). Note that the third main word order possibility, i.e. V1, does not provide enough data for a comparison across social rank and gender within the regions, apart from the fact that there is much less variation in V1 clauses, which generally prefer single negation (89% of cases; Table 4 in Section 4). It should also be noted that the lexical-semantics and the phonetic environment did not show any influence on the choice of negation (Section 5 and 6), and that we therefore will not incorporate these variables in the present section. Finally, it should be noted that differences in complexity will be ignored, even though these did seem to correlate with the negation types (Section 7). Incorporating the verbal complexity value and the proximity value, however, would in many cases diminish the data to values too low for comparison. In the present approach, some cells already have to remain empty or have only low values. We therefore decided to exclude what seems to be the least important variable of those that influence the choice of the negation type. In short, we will investigate main clauses and subordinate clauses in Amsterdam and Zeeland across social rank and gender, in order to reach as detailed a picture as possible of the influence of these social variables on the various negation types.

The overall proportion of single negation in Amsterdam was 67% (Table 3, Section 3). Figure 4 displays the proportion of single negation in V2 main clauses (117 tokens) and V-final subordinate clauses (115 tokens) in letters linked to Amsterdam, divided per social class. Note that there are no data for subclauses in the LC, as there are no negated subclauses in the LC letters linked to Amsterdam.

Figure 4 shows that generally single negation is more often used in letters from the upper ranks than in letters from the lower ranks. The trend is clear in both constructional contexts (although the LC main clause score seems to deviate), and even more obvious in the case of subordinate clauses, where the scores range from 33% in the LMC to 100% in the UC. Recall that subordinate clauses in general constitute the most conservative context (Table 4, Section 4). This means that at a time when single negation has become the preferred variant overall, the highest degree of variation is found in subordinate clauses. Figure 4 suggests that this relatively high degree of variation, with still a considerable number of bipartite negations being used, generates more social variation than a constructional context does where single negation has clearly become the dominant variant. Figure 4 also shows that the change is almost completed in the Amsterdam UC, where single negation is used in well over 80% of the tokens in main clauses as well as subordinate clauses.
The gender effect established in Section 8 is borne out by the results for Amsterdam in Figure 5. In both main clauses and subordinate clauses, the proportion of single negation is lower in letters written by women than in letters written by men. The difference is again most clear-cut in the case of subclauses, where the proportion of single negation drops to less than 50% in letters by women. Male writers, on the contrary, produce
approximately 80% single negation in both constructional contexts, indicating that the change is nearing completion.

Having established clear social and gender patterns in Amsterdam, we turn southward and focus on Zeeland, where the overall proportion of single negation was much lower, viz. 52% (Table 3, Section 3). In Figure 6, we present the proportion of single negation in V2 main clauses (124 tokens) and V-final subordinate clauses (127 tokens) in letters linked to Zeeland, divided per social class. Note that there are no data for the LC.

Figure 6 shows that there is considerable social variation in letters linked to Zeeland. In the LMC, single negation reaches only 20 to 30%, while the change is virtually completed in the UC. Interestingly, the UC pattern in Zeeland is in fact very similar to the UC pattern in Amsterdam (Figure 4), suggesting that as the change had progressed much more in Amsterdam, UC writers from Zeeland converged to the Amsterdam pattern, while bipartite negation was probably still very much present in the spoken language of Zeeland. On the assumption that bipartite negation was still fairly common in their base dialects, this would mean that UC writers created considerable morpho-syntactic differences between their spoken and written language.

Furthermore, the Amsterdam results showed the strongest effect of social variation on subordinate clauses. The Zeeland results in Figure 6, however, do not show any substantial differences between main clauses and subclauses. This may be related to the fact that the overall proportion of single negation was much lower in Zeeland. Thus there was a relatively high degree of variation in main clauses as well, which in turn gave rise to the clear social patterns as shown in Figure 6.

![Figure 6](image-url)

**Figure 6.** Single negation in main clauses and subordinate clauses in Zeeland by social class.
Turning to gender, Figure 7 presents the data for the letter writers linked to the province of Zeeland. Figure 7 shows that the low proportion of single negation in subordinate clauses is not subject to gender variation. Both male and female writers score about 40% single negation, which confirms the idea that bipartite negation must still have been well represented in the writers’ base dialects. In main clauses, surprisingly, women turn out to be more progressive than men, adopting single negation in nearly 80% of the tokens. We have no explanation for this reversed pattern. One option would be to assume that women were the leaders of the change in Zeeland, and, while we certainly cannot exclude this possibility, it seems to run counter to most of the evidence presented so far.

10. Conclusions

In the present chapter, we have identified quite a few variables, both internal and external, that influence the choice of the negation type in private letters from the seventeenth century. Our research confirms the suggestions made in earlier studies that the two most important variables are region and construction type, having found clear evidence that the change from bipartite to single negation was subject to regular regional and constructional diffusion. The change spread from the north to the south, with North Holland reaching 88% single negation in the seventeenth century, a proportion that gradually drops when moving southward, to ultimately 42% in Flanders. Our results for Holland and in particular for Amsterdam (67% single negation) and South Holland (49% single negation) differ significantly from the results of Burridge.
(1993), who established about 100% single negation in Hollandic texts from around 1650. The change also affects different construction types at a different pace, with V1 being the most progressive and V-final subordinate clauses the most conservative construction type (89% and 56% single negation respectively). We were also able to explain the difference between normal V2 main clauses and V2 main clauses with subject-verb inversion by arguing that the relatively conservative behaviour of inverted clauses was due to the historical negator fulfilling an additional function as discourse marker.

Furthermore, the pronoun men ‘one’ provided a very progressive context, with 23/24 tokens being single negations. This is probably a haplological effect on the level of the lexeme, as we were not able to show any influence of the phonetic environment in a study focusing specifically on ⟨n⟩ and ⟨en⟩ in the left context of the historical negator en. We could also not find any lexical-semantic influence of the choice of the negation type. What we did find, is that single negation was more frequent in relatively long clauses and in clauses with complex verbal clusters, pushing back bipartite negation to relatively short clauses with simplex verbs.

In Sections 8 and 9, we focused on the importance of social rank and gender on the choice of negation type. Leaving aside the surprisingly progressive behaviour of women from Zeeland (Figure 7, Section 9), we found two main effects, viz. that single negation was more frequent in the upper ranks than in the lower ranks, and more frequent among men than among women. These results are in line with earlier observations that negation had risen highly above the level of social awareness. The markedness of bipartite negation resulted in the Zeeland UC behaving similarly to the Amsterdam UC, despite the fact that bipartite negation was probably very frequent in Zeeland base dialects.
CHAPTER 11

Harvesting
Reflection and evaluation

1. Harvesting: Introduction

In the present monograph, we considered both the external variables of gender, social rank, age, time and region, and the internal variables of phonological context, formulaic context and construction type. The evidence from the confiscated private letters, examined at different linguistic levels, revealed and detailed gender and social class variation as well as language change within the time span of one hundred years. Regional variation was found in some case studies, but, remarkably, few age differences were established. In the previous chapters, we considered different explanations for the patterns of variation and change. In the present, final chapter, we will compare the results for the various phenomena, draw our conclusions (Section 2) and discuss and evaluate the explanations of community practices and politeness theory presented for separate phenomena (Section 3). In Section 4, we will focus on the writing process, and particularly the letter-writing process, itself and degrees of literacy and writing experience in the late seventeenth- and eighteenth-century communities, issues that were discussed in most of our individual case studies. Taking into consideration the full range of our results, we will present writing experience as the overarching explanation that, in our view, most comprehensively covers the patterns of variation and change revealed. Being aware of the complex nature of change, we will also briefly consider the role that internal factors appeared to play in the case of particular phenomena (Section 5). Ultimately, we will reflect on our chosen approach and evaluate the insights gained, both for the history of Dutch and in a general historical-sociolinguistic perspective (Section 6). Final conclusions will be drawn in Section 7.

2. Mapping variation and change

Having taken the tracing of variation, invisible in most printed texts, as one of our main aims, we first present an overview of the regional and social variation revealed and the diachronic differences that have become apparent.

Regional variation that undoubtedly characterised the spoken language of the past was found in the written language of our seventeenth- and eighteenth-century letters
for some of the phenomena examined. At the phonological level, traces of regional variation appear in the orthography of letters, although the degree to which this phenomenon occurs proved to be relatively low in the case of all the stable dialect features examined, i.e. $h$-dropping and $h$-prothesis, etymologically distinct long $e$’s, the Germanic $sk$ in initial position and $ft$ (Chapter 2). In comparison with supraregional spellings, localisable spellings that reflect dialect features constitute a minority. In many cases, moreover, only a few writers were responsible for a large number of the dialect spellings, which characterises dialect interference as a marginal phenomenon, linked to the practice of only a few individuals. This low degree of orality, in the sense of dialectal usage, is remarkable, when taking into account the assumed relatively oral character of private letters, and it obviously calls for an explanation, which will be given below.

In the case of forms of address (Chapter 6), hardly any regional variation was found. This is a striking feature, since regional variation is known to prevail in pronominal usage of present-day Dutch and may be expected to have been an important variable in the spoken language of the past. In the written language of our letters, however, both the pronominal forms $gij$ “you” and $u$ “you” and the nominal forms $ul$ “your love” and $ue$ “your honour” are used in all regions represented in our corpus, without significant differences. The virtually absent effect of region on the use of forms of address is a further reminder that we are dealing with a written variety of Dutch with many supraregional characteristics.

Regional variation did not play a role in the syntactic and discourse phenomenon of chaining (Chapter 7), nor in the major morphosyntactic shift from $d$- into $w$-relativisation (Chapter 8). It was, however, a factor in the case of two regionally diffused changes: morphophonological $e$-apocope (Chapter 9) and the syntactic development from bipartite into single negation (Chapter 10). In these cases, the data from our letter corpora provided a more detailed picture of the regional spread of the two ongoing changes in the seventeenth and eighteenth centuries. At the same time, the clear routes of change from north to south, that is from North Holland to Amsterdam, through to South Holland and then to Zeeland, justified our distinctions of these four regions and also showed that the metropolis of Amsterdam with its particular demography fits into the regional pattern without deviating from it.

Diachronically, the occurrence of phonological dialect features and of the older variants of final schwa and bipartite negation appears to decrease from the seventeenth to the eighteenth century, a finding that calls for some explanation. The data, both synchronically and diachronically, thus lead to the conclusion that the representation of dialectal usage is low and that the letters were strongly influenced by supraregional writing practices and standards. This conclusion has far-reaching implications for our understanding of the nature of the private letters. The private letters should not be considered as straightforwardly reflecting the spoken local dialect, but as specimens of a hybrid nature which, apart from marginal dialect features, definitely show elements of supralocal writing conventions or, more generally speaking, combine elements of speech with those of writing practice (Rutten & van der Wal 2011). An illustrative
example of the role that both speech and writing conventions may play at various stages of an ongoing change is the case of schwa apocope (Chapter 9). The data for increasing schwa apocope in the late seventeenth century had to be interpreted as a process in the spoken language of the different regions, when there was evidence of a contradiction between spoken language with progressive apocope and written language, where old writing practices showed either variation or a clear preference for schwa forms, in particular by UC men in North Holland and Amsterdam. In the eighteenth century, however, apocope had progressed so far that a new and schwa-less writing norm developed that was then adopted by UC men as a new convention (cf. Chapter 9, Section 7).

Historical-sociolinguistic analysis forms the core of this volume, and patterns of social variation were revealed in an overwhelming number of cases. Only in the very few cases of what appear to be marginal phenomena (such as sk in initial position and the ft data in Chapter 2), are the data so scarce that conclusive evidence for social class and gender distributions cannot be established, despite the fact that these analyses were initially carried out in order to gain insight into such variational aspects. In the case of formulaic language, we found clear gender differences and convincing social class patterns in that a variety of different formulae appear to be used more by women than by men and by writers from the LC and LMC more than by those from the UMC and UC. We demonstrated similar results for clause chaining by ende or en, as women and the lower ranks applied this oral feature more than men and the higher ranks (Chapter 7). We would like to draw attention to the striking finding that a similar pattern appears to occur both for a written language feature such as epistolary formulae and an oral feature such as chaining by ende/en. We will return to this issue and put forward an explanation below.

Furthermore, in the cases of new incoming variants, we also observed gender and social class variation, both synchronically and diachronically. From the established variation, we were able to conclude that men and the upper ranks clearly took the lead in the rise of the new epistolary form of address, uec, the new w-relative adverb, and single negation. These three phenomena were thus convincingly determined as changes from above, in the sense of spreading from the upper ranks. The initial variation in the case of e-apocope characterises this change as a change from below both in the sense of below the level of awareness (for most writers) and as most frequently occurring among the lower ranks (and among women) before being adopted by the higher ranks (and men) in the late eighteenth century.

Gender differences and differences between the four social classes thus appear to justify both the gender distinction and the hypothesised social stratification that we discussed in Chapter 1 (Section 5). We have to add that we only occasionally found convincing age differences such as in the case of schwa apocope in which clear generational patterns of writers below the age of 50 on the one hand and writers above the age of 50 on the other attest to a steadily proceeding change for most regions in the seventeenth century. This could point to a change initially below the level of awareness, but it should be noted that the maintaining of e-variants by upper rank men at the same
time implies awareness of this feature as a social or stylistic marker. We may conclude that the explanatory power of age with respect to language variation and change appears to be limited for the phenomena examined in our corpus, whereas gender and social class prove to be much more reliable predictors of variation and change.

Considering the fact that we are dealing with a range of changes in progress (e.g. formulae, negation, terms of address, relativisation), it is surprising that age does not prove to be an important variable in our letter corpora, whereas gender and social rank appear to be much more important. As we will argue below, this can be explained by taking into account that writing is a conventionalised practice. Mastering the conventions depends on the degree to which letter writers are part of the written culture. Whereas gender and social rank are reliable predictors of a letter writer’s education and occupation, and thus of the degree to which he or she participates in the written culture, there is not such an obvious link between age and education or occupation, in particular not when dealing with adults. In other words, when writing conventions are changing, it is first and foremost letter writers who frequently read and write who notice this. These people may be of any age, but they are more likely to be men than women, and rather part of the upper and upper middle ranks than of the lower and lower middle ranks.

For almost all phenomena that we examined, we discovered remarkable changes within the time span of one hundred years. Fewer phonological dialect features were found in the eighteenth century than in the seventeenth century; the frequency of various formulae and of clause chaining by *ende* or *en* decreased in the eighteenth century. At the same time, an increase in frequency was observed in the cases of the epistolary form of address *ue* and the *w*-relative adverb, whereas single negation was the prevailing option in the late eighteenth century. These diachronic changes also require an explanation, in the same way as the presence or absence of regional variation and the occurrence of gender and social class patterns do.

3. **Explanations: Community practices and politeness theory**

The resulting descriptions of social and regional variation and the established social patterns challenge us to provide an overarching explanation of the differences and patterns revealed. For various phenomena, previous linguistic literature refers to major explanations such as particular social group or community practices, politeness theory or education and writing experience (i.a. Laitinen & Nordlund 2012; Taavitsainen & Jucker 2003; Nevala 2004b; Vandenbussche 1999; Elspaß 2005). In previous chapters, we considered these options and critically assessed social group or community practices as an explanation in the case of epistolary formulae and politeness theory in the case of forms of address. In both cases, we concluded that these explanations might play a role in some respects (cf. Chapter 5, Section 5 and Chapter 6, Section 5), but that
they cannot explain the full range of the data and thus cannot be considered as the major, overarching explanation of our results.

If the use of formulaic language were a group-related common practice that is characteristic of the group with which writers identify themselves and independent of the level of writing experience, the differences between social groups may have been markers of group languages. Assuming that social group or community practices were determining factors for variation in the use of formulae, a number of questions would still remain, such as why the upper ranks use fewer formulae than the lower ranks, why women use more formulae than men, why the use of formulae decreases from the seventeenth to the eighteenth century and why in general converse patterns do not occur. The same questions would arise if we were to explain the differences in chaining by *ende/en* as instances of different social group practices or different styles. Although social identity roles may to some extent also come into play (cf. Chapter 5, Section 5), social group practices cannot be considered as the most convincing and insightful explanation.

In the same vein, politeness theory and its concepts of negative and positive politeness, that is frequently applied in research on forms of address in various languages, does not adequately cover our data and leaves most of our results without a convincing explanation. Although some variation might be explained by different politeness strategies (cf. Chapter 6, Section 3.1 and 3.2), we made a surprising discovery of a dominant negative politeness strategy (in the nominal forms of address *ul* “your love” and *ue* “your honour”) in private letters between spouses, siblings and friends. Furthermore, the very low proportion of the unambiguously positive polite *j*-forms was also striking, given that most of the letters were written to close family members. Finally, the diachronic increase in epistolary forms of address from the seventeenth to the eighteenth century remained an unexplained phenomenon, and their usage alongside other forms of address, which implies a supposed mixture of positive and negative politeness strategies, also raised more questions than the politeness approach answers. In other words, no satisfactory explanation was found to clarify why supposedly negatively polite forms abound in both centuries – in private correspondence.

Interpreting the forms of address primarily as conventional signs (not necessarily invoking either negatively or positively polite associations) and taking a sociolinguistic perspective, we revealed both a clear distribution of forms of address, patterning across social classes, and convincing diachronic developments. We saw epistolary forms increasing against pronominal forms such as *gij* and *j*-forms in the seventeenth century and even more strongly in the eighteenth century. The rise of *ue*, which began to replace the older form *ul* in the seventeenth-century corpus, was shown to be a gendered change from above, in that it was introduced by the upper ranks and spread from there to the lower ranks, and in that it was mostly men who led the change. In other words, we established patterns of social variation and a route of change that were similar to many others in our case studies. These strikingly similar patterns and changes can be explained by focusing on the writing process itself and by considering writing experience as the major explanation.
4. The writing process and different degrees of writing experience: Connecting all the data

4.1 Acquiring writing skills

In Chapter 1 (Section 6), discussing the autograph or non-autograph status of letters, we already touched upon degrees of literacy and of writing experience, a topic that was referred to repeatedly throughout the present monograph. Taking literacy and writing experience into account, we also formulated hypotheses, for instance, for the use of formulaic language and for the frequency of \( w \)-relative adverbs that were all borne out by the data (cf. Chapter 4, Section 2.4 and Chapter 8, Section 3). Here we will again consider the historical context of education, literacy and writing experience in the seventeenth- and eighteenth-century Northern Netherlands and will relate this background information to the presented overarching explanation.

We would like to stress once more that in the seventeenth- and eighteenth-century Dutch Republic of the Seven United Provinces the degree of literacy was relatively high, compared to various other western European countries, due to easy access to a regionally dense network of elementary schools. At the same time, we have to be aware of the various stages or degrees of literacy, varying from almost complete illiteracy, being able to read and having elementary writing skills to being fully literate in the sense of having extensive reading and writing competence. Due to the practice of teaching reading and writing successively, more people were able to read than to write and, depending on their occupations after schooling, more people were able to write only their signatures and brief notes than had more advanced writing skills and greater writing experience. What we call writing experience was operationalised with reference to literacy, schooling and the writers’ daily routine (cf. Vandenbussche 1999, 2002; Elspaß 2005: 45–46; Rutten & van der Wal 2013). In their daily routine or occupation, writers had, in other words, active or passive literacy, that is that they either regularly needed to write in their daily activities and/or occupation or had once learnt to write, but did not use their writing ability regularly (Kuijpers 1997: 514).

What literates – whether more or less experienced – had in common, is that, at some stage in their lives, they had all learnt to read and write, at school, during apprenticeships or at home, either from schoolmasters or from relatives, friends or others in their neighbourhood. Apart from private education in the highest ranks of society, reading and writing were taught at elementary schools where, according to the regulations, schoolmasters had to make handwritten models of writing for their pupils who learnt to master the technical skill of writing different styles such as gothic and cursive italic. From the diary of a seventeenth-century Dutch schoolmaster, David Beck, we know that these samples of handwriting were shown to parents as a successful means of attracting new pupils (Blaak 2009: 60–61). We may therefore assume that parents highly valued the acquisition of different hands. Convincing evidence of this assumption is found in one of the letters in our corpus, dated 14 November 1664. Writing to
her son Cornelis aboard a warship, Maria Walravens proudly mentioned the fine writing skills of his ten- and fourteen-year-old brothers. She even enclosed two samples of their writing, which give us an idea of the writing practice of copying religious phrases in different handwritings (cf. Figure 1).

Figure 1. Sample of Jacobus Walravens’ writing, enclosed in his mother’s letter.
There are a number of questions that frequently arise concerning the actual practice of letter writing. Having mastered the technical skill of writing, how did people, at a more advanced stage, learn to write a letter and how did letter writers acquire, for instance, the formulaic language that characterises our letter corpora and many other letters in contemporary European languages? Did they adopt them from the wide range of letter manuals that were frequently reprinted, or did they acquire them by some other means? Various options can be considered here: at school, the schoolmasters may have provided letter-writing models just as they did for writing different hands. At home, relatives may have trained young family members by making them copy self-written or received letters. A similar practice of copying letters may be assumed for apprentices in various professions. Considering these options, the role of the frequently reprinted letter-writing manuals and other instructive publications may be questioned.

We addressed this much-debated issue in Chapter 5 (Section 6), where we investigated the characteristics of Dutch seventeenth- and eighteenth-century private letters and compared them with the advice and models in letter-writing manuals. After having established similarities and differences between the relatively modest manuals or schoolbooks and the actual practice of writing private letters, we concluded that the differences outweigh the resemblances, to the extent that direct influence of letter-writing manuals on actual practice is unlikely. Although the manuals present letter-writing conventions and thus may have functioned as works of reference, we deemed it more likely that pupils at school and youngsters at home learnt letter-writing conventions and formulaic language, and further developed their writing skills, by participation in writing practice, that is by reading and copying letters and hearing them read aloud. Similar claims have been made for English letter writing in the long eighteenth century (cf. Austin 1973a: 13; Whyman 2009: 28–45; Brant 2006: 9–10), for German and Finnish letters from the nineteenth century (Elspaß 2005: 194–195; Laitinen & Nordlund 2012: 72–75), as well as for Dutch elite correspondence from around 1800 (Ruberg 2005: 103–114). This means that letter writing was probably acquired in a similar manner to other skills: by imitation in practice, at home, at school and at work (cf. Bannet 2005: 94).

Just as happened elsewhere in western Europe, in the Netherlands literacy in its various degrees spread at a different pace across social rank and gender, with people higher up the social ladder and men acquiring reading and writing skills generally earlier than people from the lower ranks and women (Nevalainen & Raumolin-Brunberg 2003: 40–43; Frijhoff & Spies 1999: 237–238). Literacy rates were thus generally related to social variables such as rank and gender, but were also determined by daily activities and occupation, and, in particular, the extent to which reading and writing were important in everyday life. Taking into account the socially stratified and gender-bound education differences and degrees of literacy and writing experience established by historians, we argued that these exactly parallel our gender and social class patterns of linguistic variation.
Acquired writing conventions or awareness of supraregional standards characteristic of writing experience appeared to play a major role in the case of phonological phenomena, forms of address, schwa-apocope and negation (cf. Section 2 and Chapters 2, 6, 9 and 10). The same applies to the other phenomena of formulae (Chapters 3, 4, 5), clause chaining (Chapter 7) and \( w \)-relative adverbs (Chapter 8). To fully understand these correlations, we have to concentrate on the functions and nature of these phenomena.

### 4.2 Phenomena and the explanation of writing experience

Learning to write a letter implied learning text-constitutive, intersubjective and Christian-ritual formulae. When switching from spoken to written language, letter writers could thus rely on a number of fixed, epistolary formulae. Text-constitutive formulae offered means of composing a letter along established lines, creating a text that would immediately be identified as a letter, and realising a smooth transition from one part of the letter to another. Furthermore, intersubjective and Christian-ritual formulae provided letter writers with conventionalised and generally accepted ways of verbalising information and experiences. This would have been of extreme importance for people who were not used to writing in their daily lives, and who were forced to write letters because family members or they themselves were at a geographical distance, in the country or abroad. In such cases, the use of lexicalised formulae, retrieved as whole units from memory, would have made writing easier and speeded up the writing process. If this were indeed the case, it was expected that writers experiencing problems with formulating the content of letters would prefer fixed formulae that could compensate for a lack of writing experience and compositional creativity. In other words, we expected to find more formulae in letters written by less experienced writers.

As the least experienced writers in both the seventeenth and the eighteenth centuries were women and the lower ranks, we expected these to use formulae most frequently. This assumption was borne out by the results discussed in Chapter 4, where, combining social class and gender data, we also established incremental writing experience. This implies that men score lower frequencies of formulae than women within the same social classes and that women’s scores are often similar to men’s scores one social class down the social class index. The lowest frequencies are mostly produced by the group of UC men, which typically consists of such individuals as wealthy merchants, ship-owners, academics and commissioned officers, that is, people who had a profession where extensive writing skills were needed. Increasing schooling opportunities and improved literacy rates from the seventeenth to the eighteenth century suggest a wider spread – and knowledge – of written varieties of Dutch. Diachronically, we therefore expected a levelling of writing experience and a decrease in the use of formulaic language, which was indeed attested.

The social and gender distributions of literacy, schooling opportunities, the daily contact with written varieties, in other words, of writing experience, are reflected in the social and gender distributions of the use of formulaic language. In passing, we
note that the virtual absence of regional effects on the distribution of epistolary formulae not only points at a widely shared Dutch writing practice, as mentioned above (Section 2). This phenomenon can also easily be understood from the perspective of writing experience, considering that the average literacy rates were found to show hardly any regional differences (undoubtedly due to the dense network of elementary schools). It is an explanation in terms of writing experience that covers the variation and change found not only in the use of formulaic language, characteristic for the written language of the letters, but also in the oral feature of chaining and the rise of the \textit{w}-relative adverb.

To create larger pieces of discourse, letter writers were found to use different strategies, some more oral, others prototypically written. When focusing on the presence of punctuation, the shift from more oral to more literate, referred to as the increasing textualisation of society, must have promoted the use of punctuation first and foremost among those who were most involved in the written culture. Put differently, we expected that the rise in the use of punctuation would show social and gender effects in that more men than women, and more upper-class than lower-class writers would use punctuation. This was borne out by our case study, which showed that the rise in the use of punctuation was a gendered change from above (Chapter 7, Section 5) and that the use of punctuation increased over time.

On the more oral side of the oral–literate continuum, we found semantically bleached connectives grammaticalising into discourse markers. Clause chaining \textit{en} was the most frequent of these oral strategies. It creates cohesion of two clauses and ensures continuity of the discourse from one clause to the next, but it may equally have been omitted, as it does not contribute to the propositional content. The social class and gender patterns found for chaining \textit{en} are fully covered by the writing experience explanation. Relatively more bleached connectives occurred in letters from the lower classes than in letters from the upper classes, more in letters by women than in letters by men, and more in letters from the seventeenth century than in letters from the eighteenth century. In other words, bleached connectives, as more oral characteristics of language, were found to be relatively more frequent in letters by writers who are less involved in the written culture as well as less prominent in letters by more involved writers.

The rise in the use of punctuation and the decline of clause chaining \textit{en} were gendered changes from above, which rendered the language of letters less oral and more literate. At the level of clause chaining, we concluded that the genre of private letters was less oral in the eighteenth century as compared to the seventeenth century, and that this change from oral to literate language was largely led by UC and UMC men. In summary, we are able to explain the similarity of patterns for both the written language feature of formulae and the oral feature of \textit{en}-chaining in the language use of women and lower ranks by taking into account different degrees of writing experience. The diachronic differences reflect increasing and levelling writing experience or, in other words, the increasing textualisation of society in general, which also appears to be evident in the case of the \textit{w}-relative adverb.
For the incoming variant of $w$-relative adverb, we again found clear social class and gender distributions, apart from the diachronic development towards $w$-forms, and evidence of incremental writing experience, both synchronically and diachronically, in that men were ahead of women in adopting the new form and the scores for women correspond to those of men one social class down the social class index. Both the seventeenth- and the eighteenth-century results strongly suggest that the change from $d$- to $w$-relativisation was a change from above, led by UC and UMC men. Having determined the nature of the $d$ into $w$ change, we can again apply the explanation of writing experience. In terms of syntactic complexity, the written channel promotes hypotaxis more than the oral channel does, and the rising $w$-relativisers mark hypotaxis more strongly than the $d$-relativisers (cf. Chapter 8, Section 3). Against the background of the syntacticising grammar of written Dutch, we thus understand that people who were more involved in the written culture and who, in other words, had more writing experience, adopted the incoming $w$-variant at the highest rate.

The case of relativisation thus allowed us to extend the idea of writing experience to other grammatical domains, which are less obviously related to the influence of writing, literacy and occupation. Unlike epistolary formulae, relative clauses are not prototypical elements of a written text type, and they are certainly not restricted to written language. But there is variation in how relative clauses are formed, and this variation appears to be linked to writing experience. Those who write the most use the more written-like $w$-forms, those who have less experience in writing, prefer the $d$-forms. We therefore conclude that the textualisation of society and the increasing literacy rates had an undeniable influence on the language. We witnessed the decrease in local spellings in Chapter 2, indicating supraregional convergence, which implies deviation from the spoken language. We also found a development from more oral to more literate language in the syntactic case of $en$-chaining. In the case of the $w$-relative adverb, we encountered such a deviation from the spoken language towards more written language, in the field of morphosyntax. A further result exemplifies avoiding oral characteristics: the established very low proportion of the most informal $j$-forms of address which we would expect to occur frequently in letters to family members and close friends.

The $j$-forms of address, which must have been common in the spoken language, appeared to be marginal in our corpora. In Section 3 above, we dismissed the interpretation of our results as phenomena of either positive or negative politeness and considered the forms of address as conventional signs which may play a role in supraregional writing conventions. In such an interpretation, $j$-forms are considered as traces of the spoken language. Clearly following writing conventions, letter writers occasionally diverged from these conventions, admitting oral features into their writing. The use of pronominal $gij$ and $u$ as forms of address bears witness to the fact that the writers aimed to use a supraregional variety, although obviously with occasional interference of local features such as $j$-forms. The very low proportion of $j$-forms illustrates the strong conventional usage. At the same time, the tendencies that were
identified in the use of $j$-forms reflect the different degrees of writing experience and varying awareness of prevailing writing conventions. Just as in previous cases, women and lower ranks appeared more likely to use the oral $j$-forms than men and upper ranks. Interestingly, the same applied to people from abroad, who were assumed to be less familiar with the current writing conventions: they appeared to use the oral $j$-forms more often than writers born and raised in the Netherlands.

What the present monograph clearly illustrates is that writing experience proves to be the convincing major and overarching explanation of the patterns of variation and change that we revealed in our letter corpora. Being aware of the complex nature of change, we also considered factors other than social and regional ones in the case of particular phenomena, which led to additional results and shed light on other possible influences.

5. The complex nature of change

The complex nature of change became clear in the case of negation, when various possible influences on the occurrence of single negation were examined, such as the impact of a semantic-lexical and a phonetic or phonological context. Except in the case of the lexeme $men$, it was proved that there was no impact from these contexts. The influences of region and of construction type, however, were confirmed and further detailed. Furthermore, a newly examined factor of complexity was shown to play a role: relatively long clauses and clauses with complex verbal clusters did appear to influence the occurrence of single negation, relegating bipartite negation to relatively short clauses with simple verbs. The results from our letter corpora clearly showed that the change from bipartite to single negation should be interpreted primarily as a regional change. This change, initially below the level of awareness, gradually became a phenomenon above the level of awareness and was commented upon in grammars and other manuals. That negation had risen highly above the level of social awareness corresponds with the higher frequency of single negation in the upper ranks and among men.

In the case of schwa apocope, we identified conservative and progressive left phonological contexts, while arguing that the influence of the right context was less important. Although apocope of final schwa may have been promoted by hiatus avoidance in spoken language, in written language this effect appears to be only modest. We were also able to show that both lexical and constructional diffusion play a role. Different verbs exhibit very different proportions of zero endings, suggesting that lexical diffusion is an important factor. Although a formulaic context as such does not promote the presence of final schwa, interestingly, we found evidence that a particular verb may behave differently depending on the formulaic or non-formulaic context in which it is used. This is proof that constructional diffusion is at stake. Finally, the zero subject construction appeared to provide a schwa maintaining context, a result that holds across all external variables, indicating that this particular convention was well-known to most letter writers in the western Low Countries.
In the case of the \(w\)-relative adverb, we again established, just as in the case of schwa apocope, that a formulaic context as such does not necessarily promote the maintaining of the old variant. While some formulae were conservative, preferring \(d\)-forms, others were progressive. Formulae adopted the incoming \(w\)-relativisers at different paces, confirming that language change proceeds via constructions, and that we can therefore speak of constructional diffusion. Furthermore, we found that, generally speaking, embedded relative clauses were most conservative, and that appositive relative clauses in final position, and continuative relative clauses in particular, were most progressive. This we interpreted within the syntacticisation framework: the relatively low integration of appositives into the main clause is repaired by adopting a syntactic strategy that explicitly codes a pragmatic function (by the \(w\)-form), viz. the function of keeping the discourse going.

6. The fruits of confiscated letters

The historical-sociolinguistic approach adopted in examining the confiscated letters has borne rich fruits. First of all, we gained important insights into letter-writing practice in late-seventeenth- and late-eighteenth-century Dutch societies. This letter-writing practice showed clear social class and gender patterns for both periods. For various linguistic phenomena, our exploration revealed the different language use of men and women from different social classes. Whereas the seventeenth-century printed language still allowed a fair degree of variation, the written language of our late-seventeenth-century letters exhibited an even higher degree of both social and regional variation. Moreover, alongside the considerable uniformity of eighteenth-century printed language, a still impressive degree of variation was revealed in late-eighteenth-century letters. What we thus found is not a uniform writing language, but a multifaceted picture of diversity that replaces the traditional view of limited variation for the seventeenth century and considerable uniformity for the eighteenth century.

For our exploration, we selected phenomena characteristic of the text type of letters (formulae, forms of address), remarkably frequent phenomena (formulae, forms of address, clause chaining) and familiar changes during the centuries involved (negation, schwa apocope and \(d\) into \(w\)-relative adverb). We also wanted to examine phenomena at different linguistic levels: phonological (sounds and signs), lexical (formulae) and morphosyntactic (all other phenomena). This selection appeared to imply many changes that could be considered changes from above, that is both above the level of awareness and spreading from the higher social ranks. From these remarkable results, at first sight we should not conclude that in language use changes from above are more frequent than changes from below. What we should realise and what these results also indicate, however, is the respect in which written language differs from spoken language. Whereas spoken language frequently exhibits changes from below, written language is the reflection of a more regulated practice with particular writing conventions. Letter
writers switch into, in Austin’s words, a more “formal mode of thinking” and become more aware of prevailing conventions (Austin 1994: 285). In the seventeenth and eighteenth centuries, these conventions largely appear to be determined from above.

This observation ties in with other results that provided insight into the hybrid nature of private letters. Contrary to the traditional picture based on mainly printed seventeenth- and eighteenth-century sources, our private letters show, at first sight, an impressive degree of variation at various linguistic levels. Dialectal phonological variation, however, occurred only to a fairly limited extent. We did find unambiguous traces of the spoken language, which renders the letters valuable for historical-phonological and dialectological research, but rather than local spelling, we frequently found supralocal orthographical levelling. Without getting caught in the trap of the standard language ideology, it should be concluded that in most cases, and in most letters, writers converged towards supralocal writing practices, and they did so to a far greater extent than they represented local features in their writing. This proves to be an important insight.

Putting pen to paper and switching from spoken to written language involved supralocal conventions and employing an intended supraregional variety, in the case of all social classes and for men and women alike. In these private letters, both supraregional writing conventions and changes from above were thus found which ultimately influenced the lower- and middle-class writers. Returning to the previous discussions of the language of immediacy, the language from below approach and the hybridity of private letters, we conclude that we did find more variation in our letter corpora than in contemporary printed publications. At the same time, we found relatively little influence from spoken language. Therefore, it is important to stress once more that what we traced in the confiscated private letters is to a large extent not the everyday spoken language of the past, but rather another linguistic variety: the contemporary informal written language. The hybrid nature of this informal written language with respect to the oral–literate continuum (and the concepts of immediacy and distance) became clear in the previous chapters of the present monograph. At the same time, this variety represents language from below, that is the informal written language of letter writers from below the ranks of the ruling class and nobility.

Not only were the diversity and patterns of variation revealed, the overarching explanation for writing experience and different degrees of literacy was also provided. We have thus come to a better understanding of what happens when people start to write and communicate with distant addressees. For particular phenomena, we were able to prove that this explanation of writing experience is more convincing than other possible explanations. For instance, interpreting forms of address as conventional signs and taking a sociolinguistic perspective was demonstrated to cover the data to a larger extent than could be done by politeness theory.

Distinguishing between autographs and non-autographs was a major methodological issue in our research. After having established the autograph status of letters, we were able to carry out our historical-sociolinguistic analyses for the variables gender, rank and age. When examining formulaic language, we compared the results from autograph letters to those of non-autograph letters, written by social or professional
writers. The striking frequency differences suggest that social and professional writers of private correspondence must have relied on formulaic language, self-evidently not out of lack of writing experience, but as convenient tools of fixed formatting which rendered communication easier. This seems to be an obvious idea. Without further evidence, however, this idea was simply a plausible hypothesis. In the present monograph (cf. Chapter 5, Section 3) we presented quantitative linguistic evidence which proves that it must have been common practice for social and professional writers to rely on formulaic language. Writing in the name of someone else triggered the use of formulaic language. Again, this gives more insight into the practice of everyday communication: the epistolary formulae function both as a safe haven for writers with little writing experience and as a convenient tool of fixed formatting for social or professional writers.

Exploring the confiscated letters, we thus achieved a range of important insights. During our research, we were frequently asked by fellow linguists whether we also encountered new and unexpected phenomena, in particular for the history of Dutch. We did indeed make a number of surprising and unexpected discoveries, such as a late-eighteenth-century Papiamentu letter that appeared to be the third oldest Papiamentu text ever to have been discovered (cf. Jacobs & van der Wal, to appear). This and other letters with non-Dutch fragments reveal multilingual situations and code switching mainly by letter writers living abroad. Apart from surprising views on the cultural-historical and linguistic context, we also found data that allowed us to fill gaps in the history of Dutch. Generally speaking, this occurred when presenting the linguistic data related to the four social classes distinguished. These data all largely complement the traditional picture of a language history based primarily on the highest social rank of nobility and the ruling class. The same applies to the language use of female letter writers, that was previously largely unknown. The letters also allowed us to discover unknown orthographic writing practices, to detail particular regional variation and to establish routes of linguistic change. We mention a few examples to illustrate this. For orthographic writing practices, we confirmed the existence and use of the lesser known variant ⟨ee⟩ alongside the commonly used variants ⟨ae⟩ and ⟨aa⟩ in open syllable (Chapter 2). For regional variation, the letters gave new evidence for the spread of h-dropping and h-prothesis in the region of North Holland (Chapter 2), the regional spread of single negation was further detailed (Chapter 10) and its change from above was proved. In the case of the familiar change from bipartite to single negation, it became evident that the traditional view of a prevailing single negation around 1650 does not reflect linguistic reality. The same applies to the use of forms of address, the social class and gender distributions of which appear to deviate largely from the traditional view and the practice of literary authors. At the lexical level, we gained a detailed view of an abundant range of Dutch formulae.¹

¹. Furthermore, we may note that the letters also provided many new variants and various unknown lexical items that add to the lexical databases of the Leiden Institute for Dutch Lexicology (INL).
7. Conclusions

In Chapter 1 (Section 2) we mentioned four separate goals that were our aims in this monograph: mapping variation and change, presenting explanations for the variation and change found, filling gaps in the history of Dutch and contributing to the international historical-sociolinguistic debate. In the present chapter, we can establish that these goals have been achieved. Moreover, at all linguistic levels the letter corpora presented new and surprising results. These results confirm the fruitfulness of a historical-sociolinguistic approach, and show both the importance of writing experience as an overarching explanation and the value of the compiled *Letters as Loot* corpora. At the closing of our research programme, these rich corpora and the related metadata were presented and made available for linguists, historians and others interested in the extraordinary letters at http://brievenalsbuit.inl.nl.2

![Figure 2. Pen and papers in one of the HCA boxes (National Archives, Kew, UK).](image)

Overlooking the harvest, we consider the monograph’s main contribution to historical sociolinguistics to be the detailed and empirically sound assessment of the everyday language of ‘normal people’ at the crossroads of the spoken language and their simultaneous awareness of supraregional writing practices. In this respect and for the history of Dutch, we can speak of the emancipation of the major part of language users. Access to their language of the past was gained by the unique source of confiscated letters, dating from times of war, a source that continues to form a treasure trove for historical linguists.

2. The internet application http://brievenalsbuit.inl.nl was developed in close collaboration with the Leiden Institute for Dutch Lexicology (INL). We gratefully acknowledge the contributions of various INL colleagues: Marijke Mooijaart and her lexicological assistants who carried out what appeared to be an enormous task of tagging and lemmatizing the corpus and Katrien Depuydt and Jesse de Does who took care of the technology and provided an attractive, user-friendly application.
References


Austin, Frances. 1973b. *Epistolary conventions in the Clift family correspondence (Concluded)*. *English Studies* 54. 129–140. DOI: 10.1080/00138387308597547


Breton, N. 1645. *Nederduytsen briefdragher*. Amsterdam: Jan van Hilten.


*DynaSAND*, see Barbiers et al. 2006.


MAND = *Morfologische atlas van de Nederlandse dialecten*. Vol. II. Ed. by Ton Goeman, Marc van Oostendorp, Pieter van Reenen, Oele Koornwinder, Boudewijn van den Berg & Anke van Reenen. Amsterdam 2008: Amsterdam University Press.


MNW = Middelnederlandsch Woordenboek. gtb.inl.nl.


ONW = *Oudnederlands Woordenboek*. gtb.inl.nl.

Paardekooper, Piet 2006. Bloei en ondergang van onbeperkt ne/en, vooral dat bij niet-woorden. *neerlandistiek.nl* 06.02 (see http://www.neerlandistiek.nl/06.02/).


Letters as Loot


Letters as Loot


Sweetser, Eve E. 1990. From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620904


Vermaas, J.A.M. 2002. Veranderingen in de Nederlandse aanspreekvormen van de dertiende t/m de twintigste eeuw. Utrecht: LOT.

Vermaas, J.A.M. 2002. Veranderingen in de Nederlandse aanspreekvormen van de dertiende t/m de twintigste eeuw. Utrecht: LOT.


VNMW = Vroegmiddelnederlands Woordenboek. gtb.inl.nl.


WNT = Woordenboek der Nederlandsche Taal. gtb.inl.nl


Wray, Alison. 2002. Formulaic language and the lexicon. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511519772


Websites

www.brievenalsbuit.nl The Brieven als Buit/ Letters as Loot programme website, both in Dutch and in English.

Index

A
analogy 35, 38, 39, 40, 66, 70, 326
appositive relative clause 290, 313–321, 405
ars dictaminis 76
ars notaria 76
autographs vs non-autographs 173–179

B
business letters vs private letters 208–209, 244

C
capitalisation 252, 256, 272
change from above 141, 171, 228, 229, 243, 245, 250, 269, 284, 287, 296, 299, 302, 320, 395, 397, 402, 403, 405–407
change from below 5, 395, 405
clause integration 289, 313, 314, 316, 318
coherence 172, 257, 258, 266, 319, 320, 371, 372
cohesion 257, 266, 270, 402
collocation 75, 79, 116, 117, 121
common usage verbs 374
complexity 294, 295, 305, 364, 380–384, 388, 392, 403, 404
conserving effect 318, 353
constituent negation 369, 370
constructional diffusion 305, 313, 320, 353, 354, 359, 361, 370, 391, 404, 405
collocation 75, 79, 116, 117, 121
collocation 75, 79, 116, 117, 121
coherence 172, 257, 258, 266, 319, 320, 371, 372
collection 257, 266, 270, 402
collocation 75, 79, 116, 117, 121
common usage verbs 374
complexity 294, 295, 305, 364, 380–384, 388, 392, 403, 404
conserving effect 318, 353
constituent negation 369, 370
constructional diffusion 305, 313, 320, 353, 354, 359, 361, 370, 391, 404, 405
collocation 75, 79, 116, 117, 121
common usage verbs 374
complexity 294, 295, 305, 364, 380–384, 388, 392, 403, 404
conserving effect 318, 353
constituent negation 369, 370
constructional diffusion 305, 313, 320, 353, 354, 359, 361, 370, 391, 404, 405
construction type 305, 313, 364, 368–371, 377, 387, 388, 391–393, 404
continuative relative clause 311–314, 319, 320, 405
conventionalisation 87, 93, 122, 129, 133, 207, 214, 221, 222, 225, 243, 244, 350, 359, 396, 397, 401, 403, 406
conversational routine 123
D
degree of orality 8, 19–21, 23, 45, 66, 73, 249, 286, 394
discourse continuity 109, 158, 248, 257, 258, 260, 266, 320
discourse marker 257–266, 286, 373, 392, 402
developmental writing experience 169, 234, 299, 301–304, 320, 401, 403
intended superegional variety 74, 406
intonation units 248, 271
inversion, see subject-verb inversion

The study of letter writing is at the heart of the historical-sociolinguistic enterprise. Private letters, in particular, offer an unprecedented view on language history. This book presents an in-depth study of the language of letters focussing on a unique collection of Dutch private letters from the seventeenth and eighteenth centuries, which comprises letters from the lower, middle and upper ranks, written by men as well as women.

The book discusses the key issues of formulaic language and the degree of orality of private letters, it questions the importance of letter-writing manuals, and reveals remarkable patterns of social, regional and gender variation in a wide range of linguistic features. Arguing for writing experience as an important factor in historical linguistics generally, the book offers numerous new perspectives on the history of Dutch.

The monograph is of interest to a wide readership, ranging from scholars of historical linguistics, sociolinguistics, Germanic linguistics, sociology and social history to (advanced) graduate and postgraduate students in courses on language variation and change.

“This book makes an original and worthwhile contribution to the study of the development of language in a social and historical context. The analyses are undertaken with exemplary thoroughness, presented with copious detail and compared with studies in other European languages.”

Martin Durrell, Manchester

“Drawing on unique source material and presenting detailed case studies of a variety of features from seventeenth- and eighteenth-century Dutch, this well-written and important study of private letters from writers of all social classes and different geographical backgrounds offers an entirely new perspective on the history of Dutch.”

Ingrid Tieken-Boon van Ostade, Leiden