

Vrije Competitie GW: *Language Dynamics in the Dutch Golden Age*
Marjo van Koppen & Feike Dietz (Universiteit Utrecht, UiL-OTS en ICON)

1. Project Information

Language Dynamics in the Dutch Golden Age: linguistic and social-cultural aspects of intra-author variation

2. Summary

(240 words)

The vibrant political, religious and cultural atmosphere of the Dutch Golden Age interacted with language. 17th century Dutch was a mixture of fading linguistic properties from the preceding language phase, Middle Dutch, and upcoming new ways to construct words and sentences. These language innovations were partly driven by migration, literary innovations and standardization processes.

Within these language dynamics we observe a type of language variation that has rarely been addressed before: variation within individual language users (*intra-author variation*). The famous author P.C. Hooft, for instance, uses the Middle Dutch way to express negation as well as a modern alternative. How can we account for this variation, seemingly randomly displayed by authors? This project will chart and explain the grammatical properties of intra-author variation, as well as the social- and literary-cultural factors that influenced the way individual authors used their variation in a strategic and/or creative way. The central hypothesis of the project is that the (internal) grammars of authors created a particular range of variation, which was systematically used by authors, based on contextual factors.

We develop a new line of interdisciplinary research as a necessary condition for an in-depth understanding of language variation, combining approaches from theoretical linguistics, historical sociolinguistics, computational linguistics and literary studies. We qualitatively investigate 1) how variation follows from the (internal) grammar, and 2) is related to the social and literary context, and we quantitatively investigate 3) variation patterns within and between authors and genres.

9a. Description of the Proposed Project

(2449 words)

I Scientific Quality

1. Research Question, Hypothesis, Aims

The Dutch Golden Age was a dynamic period with innovations in many domains: economy and trade, arts and sciences, politics and religion. This vibrant atmosphere interacted with a dynamic linguistic situation. The Dutch vernacular, developed into the *Lingua Franca* of the newly formed Dutch Republic, was influenced by several intra- and extra-linguistic factors (Sijts&Willemijns 2009; Sijts 2004; Wal&Bree 2008; Wal 1995). Dutch was a melting pot of fading linguistic properties from the preceding language phase Middle Dutch (like case marking), and upcoming new ways to construct words and sentences (like the more wide-spread use of prepositional constructions). Moreover, language development was spurred by large scale migration, the use of Dutch in new social domains (religion, science, etc.), the rise of new literary genres and active language regulation, e.g. via the formal agreements reached during the States translation of the Bible.

As is expected in such a vibrant situation (Thráinsson 2012 and Nevalainen&Raumolin-Brunberg 2012), all of this resulted in a wealth of language variation. This project uses this very suitable case to break new ground in the study of language variation, by combining insights of three disciplines that investigate language variation from different perspectives (theoretical linguistics, historical sociolinguistics and historical literary studies), and by profiting from current digital developments. While theoretical linguists investigate the grammatical properties of language variation, explaining how they result from the language user's mental grammar (i.e. his/her internal language system), and historical sociolinguists and literary historians focus on the way variation is used in relation to the social/literary-cultural background, we combine these fields to investigate a subtype of variation that has rarely been addressed: *intra-author variation*, the language variation within individual authors.

Using this new interdisciplinary research line, this project will explain the variation between and within 'registers' of individual language users, i.e. their 'subgrammars' in different language contexts (Mesthrie 2013:70). It will shed light on the grammatical properties combined with the full range of social-cultural factors and literary strategies influencing the variation in and between registers of 17th century authors.

General Research Question
How can we account for intra-author variation in 17th-century Dutch?

Hypothesis

The patterns of intra-author variation in 17th-century Dutch result from the interaction between the internal language system of a language user and the social/literary-cultural contexts the language is used in: the internal language system of authors makes available a particular range of variation, which is then systematically and pattern-wise used by authors in different contexts.

Aims

Central scholarly aim: charting and explaining the grammatical properties of intra-author variation in combination with the social- and literary-cultural factors that influenced the way individual authors used their variation in a systematic and strategic way.

Methodological aim: developing and enhancing a new line of interdisciplinary research.

2. Operationalization

In order to explain how we will make new steps in the research on intra-author variation, let us consider a specific case study: negation. Earlier research into negation mainly focused on diachronic changes (Zeijlstra 2004, Kemenade 2000), on describing its variation at the various language stages (Horst 2008) or on the variation between social groups (Nobels 2013). We will take a different viewpoint and analyze how and why one author uses different options in the same period: the famous writer and politician P.C. Hooft, for instance, uses in his letter to his beloved Eleonora Hellemans both the Middle Dutch way to express negation (a combination of the negative clitic *en* and a negative particle *niet*; compare French *ne...pas*) and a modern alternative (single negation: *niet*).

In order to fully understand the system behind this specific instance of intra-author variation we must take into account both the grammatical properties of the variation as well as the literary-cultural context it is used in. The internal language system of Hooft seems to systematically deploy this variation: negative concord, the unmarked way to express negation in Middle Dutch (Hoeksema 1997), has been reinterpreted in Hooft's language system as a marked, contrastive form of negation, resembling emphatic negation (Wouden 1994). Hooft for instance uses negative concord in sentences of the type 'not this, but that':

Moedwillighlijk nochtans en heb ick mij in deze diepten van droefheit niet,
Intentionally however NEG have I me in these depths of sadness NOT

geworpen maer [...] allerlei onderhoudt gezocht om mijne zinnen te verleijden.
thrown but diverse entertainment searched for my senses to divert

“However, I have intentionally **not** thrown myself in the depths of sadness, **but** I have searched to divert my attention by seeking [...] several types of entertainment.” (Hooft 1627)

The literary contexts also seems to be relevant for his choice of negation: we find more negative concord in his letter to Hellemans in which he expresses regret over her rejection, than in some of his other letters. We can therefore only fully understand the patterns of intra-author variation in 17th-century Dutch if we investigate quantitatively and qualitatively the grammars of authors in combination with their social/literary-cultural context. That's why our main research question is operationalized into the following subquestions; each of which will be answered in a particular subproject, using its own methodology (see section IV).

Subquestions & Approach		
Subproject 1	Subproject 2	Subproject 3
How does intra-author variation in 17th-century Dutch follow from the properties of the internal language system?	How does intra-author variation in 17th-century Dutch follow from and feed into the literary-social context?	Which patterns, frequencies and distribution of (morpho)syntactic variation within and between authors, registers and genres can be uncovered using digital tools?
↓	↓	↓
Hypothetical-deductive theoretical linguistic approach	Qualitative text analysis, using methods of literary/cultural studies and historical sociolinguistics	Quantitative corpus linguistics approach

Corpus Selection

We will investigate these questions through a digital corpus consisting of texts from around ten authors who were each active in several genres. The selection of authors varies along the following axes that are known to have had an impact on 17th-century Dutch (Sijts 2004):

- **Progressive/conservative use of language:** innovators who are active in standardizing and embellishing language (Hooft, Bredero, Coornhert, Vondel) as well as authors that mainly use written language as a means of communication (Michiel de Ruyter, Koelmans 2001) or are known as more conservative and non-creative language users (Johan de Brune de Oude and Johan van Beverwijck, cf. Wal 1990, Horst&Wal 1979).
- **Geographic circumstances:** authors born and raised in the Dutch Republic as well as migrants (Bredero, Vondel); authors with an urban (Hooft) and rural background (De Brune).
- **Period:** authors active in the first (Bredero) or last (Samuel van Hoogstraten) part of the 17th century, or during a long period (Hooft)
- **Genre:** per author as many genres as possible (Van Hoogstraten's novels, pamphlets, tracts).
- **Audience:** texts written for different audiences (Aernout van Overbeke, cf. Barend&Gelderblom 1998).
- **Gender:** male and female authors (Isabella de Moerlooze).

Morphosyntactic features

We particularly focus on (morpho)syntactic features that are known to display a large amount of variation in 17th-century Dutch (Horst 2008). Our main attention will be on *negation*, *case* and *infinitival clauses*. These features are of specific relevance because they underwent natural changes (i.e. changes that are also found in other languages and follow a regular (linguistically motivated) patterning of change, Gelderen 2009), and/or were the result of prescriptive/cultural pressure:

Negation	The negative cycle is a natural diachronic change found across languages (Jespersen 1917; Zeijlstra 2004), reinforced by social-cultural pressure (Wouden 2007; Paardekooper 2006).
Case, adpositions, word order	The loss of the case system (leading to more fixed word order/more adpositions) is a natural diachronic change found across languages (Kemenade 1987, Neeleman&Weerman 1999), but was hindered by language prescriptivists who tried to retain the case system (Sijs 2004, Weerman et al. 2013, Hendriks 2012).
Infinitival clauses	The extensive use of infinitival clauses is a literary (instead of natural) language innovation, mimicking the classic languages (Sijs&Willemijns 2009).

3. Scientific Embedding and Relevance

Aiming to uncover the principles and parameters of the internal language system (Chomsky 1995 *et seq.*), **theoretical linguists** extensively studied the grammatical properties of variation between languages in the past decades (Biberauer 2008). From this perspective, they have also focused on variation within individual language users, trying to establish the locus of this type of variation within the internal language system (Kroch 1989,2001; Yang 2000,2010; Biberauer&Richards 2006; Roeper 2011; Thráinsson 2012). Since theoretical linguists normally investigate speakers rather than authors, this project, on intra-author variation, introduces a new dimension to this research field.

Historical sociolinguistics – which examines the relation between language and social variables (e.g. gender, class) – also investigates intra-author variation, focusing on how the social context (e.g. text type, intended audience and/or social context of the author) interacts with variation (Auer 2015, Tieken 2013, Nevalainen&Raumoulin-Brunberg 2003,2012). Our project combines these types of research, and profits from current developments in eScience by incorporating the digital possibilities **computational linguistics** has to offer (in particular data-mining of large digital text corpora).

Early-modern literary studies have rarely studied intra-author variation in the past decades, but have developed several perspectives and concepts which could enhance our understanding of intra-author variation. From the 1960s onwards, this field developed research into rhetoric, which was mainly concentrated on the properties of a certain genre (Witstein 1969; Spies 1999; Jansen 2008). Within this tradition, some scholars analyzed the language use and style of individual authors – yet focusing on constants rather than variation within one author (Damsteegt 1978,1981; Gelderblom 1989; Jansen 2011).

By combining the theoretical and methodological expertise from these four domains, our research aims to understand the (interplay and effects of) the grammatical and social/literary-cultural influences on intra-author variation. By including literary studies in this project we will gain a more fine-grained view of (i) literary-cultural aspects influencing intra-author variation (e.g. the literary position of authors, genre conventions) and (ii) the effects and creative aspects of text(-parts). While current historical linguistics

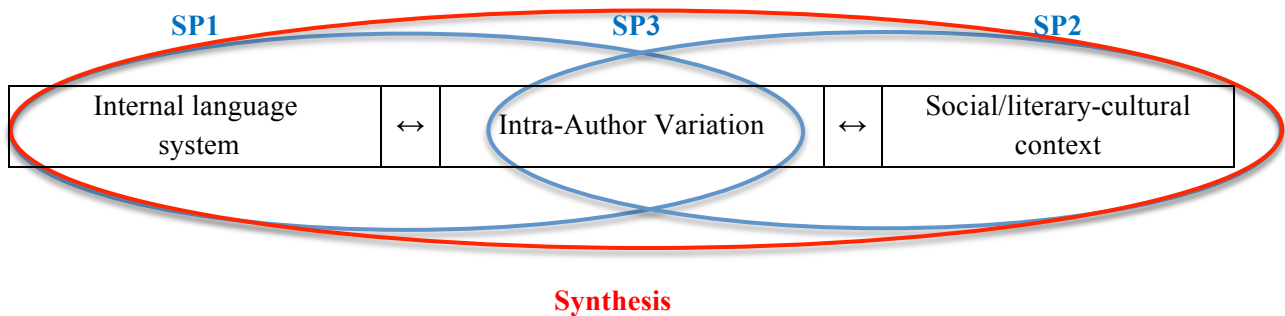
mainly restricts itself to ‘speech-like’ genres such as letters and diaries (Auer 2015, Goss&Howell 2006), this interdisciplinary project is able to include higher literary genres as well.

Using 17th-century Dutch as a fruitful new test case to understand intra-author variation (see also Overdiep 1947-48), this project will extend and expand the ongoing linguistic research into 17th-century Dutch, which is now mainly concerned with (the effects of) language standardization (Wal&Bree 2008; Sijts 2004; Goss 2002), with the description of variation in general (Horst 2008) and with sociolinguistics properties of *groups of language users* (Nobels 2013, Wal&Rutten 2013,2014).

Furthermore, this project contributes to early modern literary studies. From the 1980s onwards, literary historians turned their attention towards the cultural-historical background of authors and texts. This project feeds the ongoing research on textual and media strategies that fueled social debates (Frijhoff&Spies 1999), by analyzing the way individual authors actively and creatively contributed to the development of new Dutch idioms to discuss for instance religious and scientific topics that have long been part of a Latin textual culture. Moreover, this project sparks the expanding research on early modern authorship, by focusing on the variation within the language of one author. Literary historians have a long tradition of focusing on the strategic way early modern authors used their language (Greenblatt 1980; Geerdink 2012; Schenkeveld 2009), but mainly focused on the way in which authors are part of bigger social-cultural discourses and not so much on the variation within these authors’ works.

II Project Criteria/Design

The scheme in paragraph I.2 depicts the project design. The *first subproject* investigates the grammatical properties of the variation within and between registers and aims to explain how they follow from the internal language system. The *second subproject* explores the characteristics of variation between and within different registers, considering both contextual factors such as audiences and education as well as literary strategies and purposes. Together, the subprojects will provide a more fine-grained picture of the characteristics and functions of registers and their variation, and the way authors used and extended the variation options available in their language. *Subproject 3* analyzes the variation patterns quantitatively. In particular, this project investigates which patterns, frequencies and distribution of (morpho)syntactic variation within and between authors and genres can be uncovered using digital tools. This subproject tests the analyses developed in subprojects 1+2, as well as enhances them by detecting additional variation patterns. In *synthetic* publications, these levels and perspectives will be connected to establish the full range of factors leading to intra-author variation as well as the interplay between them.



The applicants are experts in the fields of variation linguistics and early-modern literature and are familiar with interdisciplinary research (see CV). Institutionally the project will profit from the UU-institutes for Cultural Inquiry (ICON) and Linguistics (UiL-OTS: a center that is renowned for research into (morpho)syntax, language variation, theoretical, historical and computational linguistics as well as sociolinguistics), and the Digital Humanities Lab. Members of the advisory board will regularly be consulted and will attend experts meetings. There will be close collaboration with Nederlab and CLARIAH to develop the digital components of this project (see subproject 3).

III Knowledge Utilization and Transfer

A. Publications in popularizing media

Our modern society experiences similar language dynamics as the 17th century, with the rise of new text genres, for instance *whatsapp*-messaging, influenced by both internal and external linguistic factors (age, social scenes). Concerns about modern developments lead to debates on whether language use in these text types has an effect on literacy (Spitzer 2013, Hulshof 2013). With the knowledge that we gain from language variation within the Dutch Golden age, we will partake in these debates via publications in journals targeting bigger audiences, like *Onze Taal* and *Vaktaal*. We will show how the dynamic interaction between language regulation and development leads to a more creative and richer use of the language.

B. Educative module

Anticipating the demands from educators and didactic theorists for teaching materials that invite pupils to examine language variation (Lefebvre 2014) and historical Dutch (Robben 2010), and that cover both linguistics and historical literature (Maas 2008), we will develop a digital integrated language and literature module for secondary schools (last years havo/VWO), introducing pupils to the language dynamics of the Dutch Golden Age. On the basis of acquired knowledge on how variation can be used creatively, pupils are challenged to investigate language variation of a new text type (like *chats*). This module will not only educate them in the language, literature and history of the Dutch Golden age, but they will also discover the systematic properties of language, developing their language awareness and knowledge on language as a tool in a social/literary-cultural context.

C. Scientific valorization

This project is relevant for other humanities scholars, like historians, because we unlock our cultural heritage by digitalizing texts, enriching them with various linguistic annotations, and developing tools to search through them (see below in subproject 3).

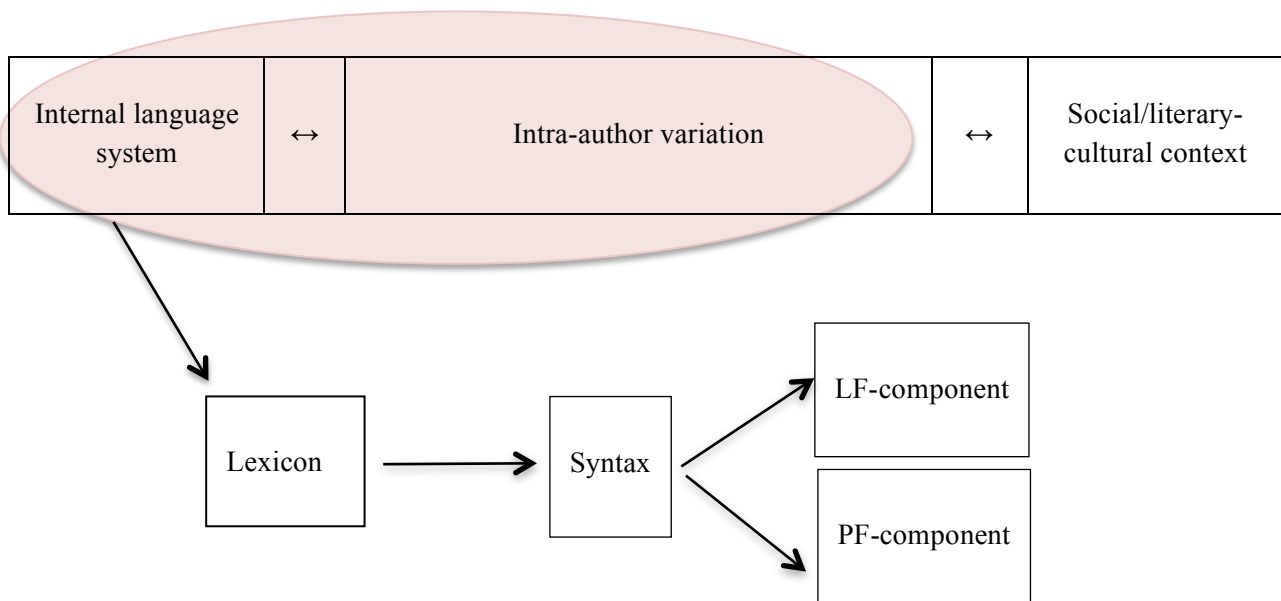
IV Subprojects and Synthesis

Subproject 1: *The relation between intra-author variation and the internal language system in the Dutch Golden Age*

(788 words)

Questions, Goals

This subproject aims to investigate the relation between intra-author variation and the internal language system, which we assume (somewhat simplified) consists of (i) a pre-syntactic lexicon, (ii) a syntactic component and (iii) the interface components PF (spelling out syntactic structures) and LF (the interface between syntax and semantics) (Chomsky 1995).



Research Question

How does *intra-author variation* in 17th-century Dutch follow from the properties of the internal language system?

Subquestions

Empirical subquestion

What are the (grammatical properties of the) variation patterns within and between the different registers of individual authors?

Theoretical subquestion

What is the locus of this *intra-author variation* within and between different registers in the internal language system?

Approach

This subproject is framed within the Minimalist Program (Chomsky 1995,2000 *et seq.*). The generally accepted hypothesis within this framework is that variation can be reduced to properties of functional items (e.g. complementizers, auxiliaries, pronouns) in the pre-syntactic lexicon and the way the PF-component spells out syntactic structures (Kayne 2005; Borer 1984,2005; Gianollo et al. 2008). We will illustrate this with our example of Hooft's use of negation in his letter to Hellemans. Our preliminary research revealed that Hooft does not use negative concord in every letter: he seems to have both a register in which there are two types of negation with their own semantic flavor, and one with only one type of negation, schematized as follows:

Hooft's internal language system

REGISTER-1

- 1:negation: niet
- 2:negation+focus: en...niet

REGISTER-2

- 1:negation: niet
- 2:negation+focus: niet

One might hypothesize that the variation within register-1 results from the pre-syntactic lexicon. In context-2, but not in context-1, it contains a focus feature. This leads to different syntactic structures as well as different spell-outs. The variation between Hooft's registers concerning context-2 might be analyzed as follows: the pre-syntactic lexicon is the same in both registers, leading to the same syntactic output structure, but the PF-component in register-1 has a separate lexical item *en* to spell out the focus feature, whereas register-2 lacks such an item and hence has to realize the structure in a different fashion (Adger 2014; Adger&Smith 2005,2010, Barbiers 2014). Another potential source of variation between registers comes from Biberauer&Richards (2006), Barbiers (2009), Roeper (2011) and Belder&Koppen (2014). They show that syntax can also be responsible for variation, providing more than one output structure which is then attributed to a certain register.

Hypotheses

1. The locus of this *intra-register* variation is restricted to the pre-syntactic lexicon and the post-syntactic PF-component.
2. The locus of *inter-register* variation is either the same as that of *intra-register variation* or it is the result of syntax producing more than one possible spell-out of which one is dedicated to a specific social/literary-cultural context.

This extension of microvariation research to include register variation within one author is an important step, as it provides us with a more complete and fine-grained picture of linguistic variation. Understanding intra-individual variation is as Schilling-Estes (2002:376) says (about intra-speaker variation, but this can be extended to intra-individual variation in general): “[...] we cannot hope to achieve a full understanding of the patterning of variation in language, or of language in general, if we do not understand its patterning within individuals’ speech as well as across groups of speakers.” Moreover, variation within one author is the ultimate example of what Kayne (2005:5) argues is the ideal experiment to investigate variation: languages that closely resemble each other make it easier to detect the abstract parameters underlying the minimal variation between them. Registers are potentially the closest variants of a language one can find. Note, furthermore, that intra-author variation provides a good testing ground to investigate how (morpho)syntactic structures interact with a sociolinguistic property like register since all other sociolinguistic variables are kept constant (e.g. age, gender, social class of the author).

Methodologically, this project qualitatively investigates a restricted number of authors, making use of the hypothetical-deductive methodology of theoretical linguistics. It focuses both on authors that use written language as a means of communication (De Ruyter, Van Beverwijck) and on literary language innovators (Hooft, Coornhert), focusing as much as possible on prose (and not on poetry) to get the best reflection of their internal language systems. We also include texts/editions written at different points in an author’s career to incorporate the dimensions of language change (Quené 2013; MacKenzie 2014; Branden 1956 on Coornhert).

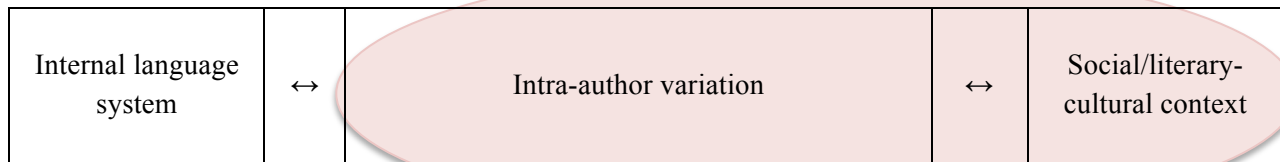
In line with the (historical) sociolinguistic literature we will start with defining register broadly as a above, a ‘subgrammars’ used in different language contexts including formal/informal language, genre, gender, social class etc. (Mesthrie 2013). Since this is still a rather crude way to define a register, we will refine our notion of register based on the input from and cooperation with subproject 2.

Subproject 2: Intra-author variation in the social/literary-cultural context of the Dutch Golden Age

(769 words)

Questions, Goals

This subproject explains intra-author variation from the perspective of the social and literary-cultural contexts in which language functioned. It aims to understand how and with what purposes authors effectively and strategically used the variation between and within the different registers which were made available by their language system.



This subproject departs from the assumption that Hooft's choice for the negative concord in his letter to his beloved female friend Eleonora Hellemans (Tricht 1955) was related to the contextual factors of text, consumption, production, and conceptions of language. In this case: the genre and private nature of the text (a handwritten love letter, based on epistolary conventions), Hooft's cultural background and networks (his high level of education; his rich network of innovative authors, Koppenol 2012), his relation to this female addressee, the purpose of his text (he was desperate to convince her) and the linguistic conventions at the time (that prescribed more modern forms of negation, Wouden 2007).

By analyzing all of these factors in a number of carefully selected cases, this subproject will shed light on how intra-author variation within and between registers was spurred on by social/literary contexts, as well as on how intra-author variation affected language used by contemporary authors. While *imitatio* was the leading principle in the early modern textual culture (Jansen 2008), we can assume Hooft's example was closely examined and either consciously followed or rejected (Tieken 1991; Auer 2006 for English).

Research Question
How does <i>intra-author variation</i> in 17th-century Dutch follow from and feed into the social and literary-cultural context?

Subquestions

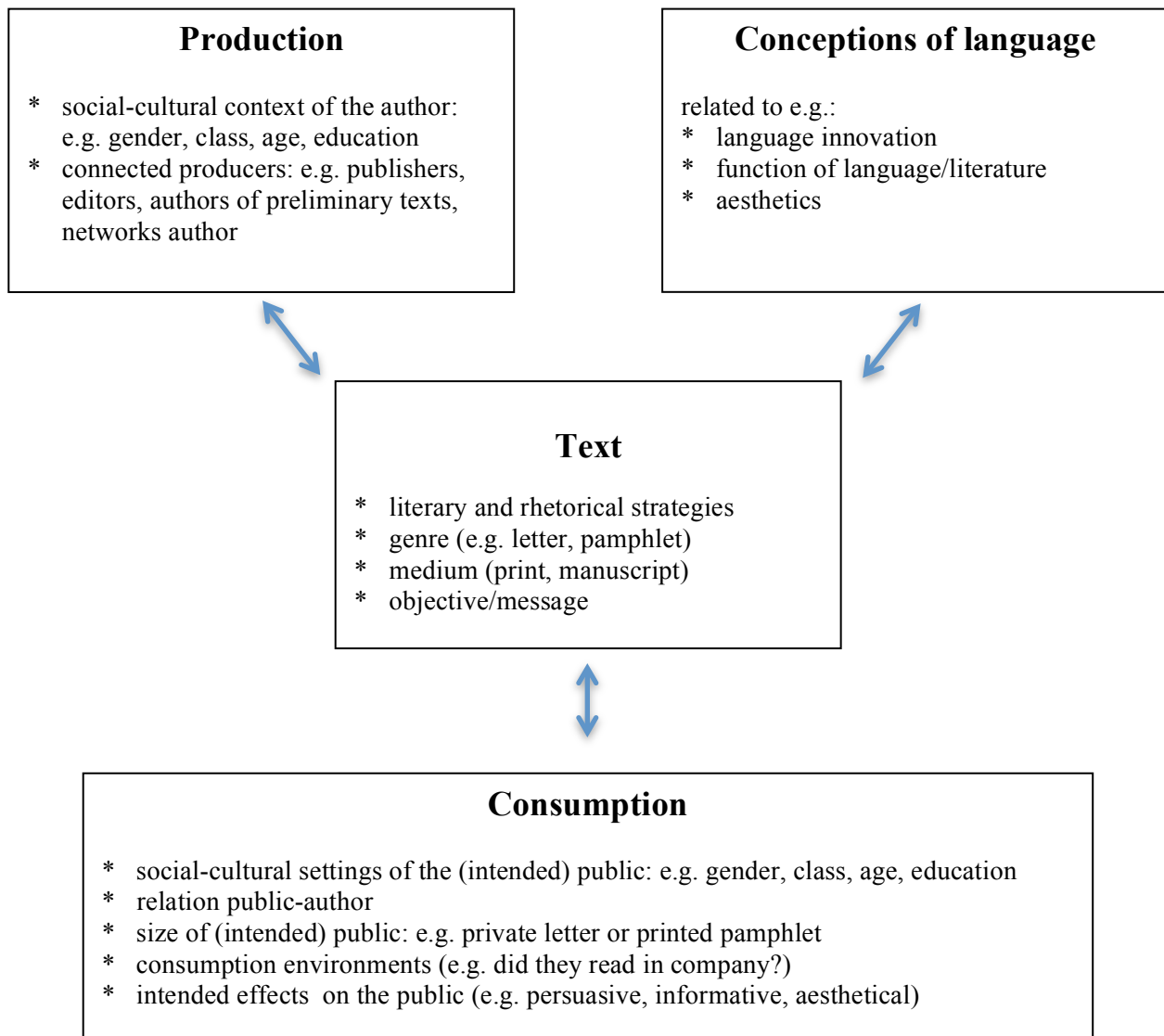
1. What was the effect of social, cultural and literary factors on the variation within and between the registers of individual authors?
2. How did individual authors systematically and strategically use the variation within and between their registers in particular social, cultural and literary contexts?

Approach

As explained in I.3, this is a new type of research question in the field of historical literary studies, and requires the development of an approach based on both historical sociolinguistic research and literary studies. Historical sociolinguistic research offers the insight that language variation was an effect of contextual factors (e.g. education of language users, genre norms) as well as a driving force behind the creative (re)shaping of these context (Auer 2015, Schilling-Estes 2002). Following this sociolinguistic approach, this subproject examines both the social context of the involved consumers and producers of texts, as well as their opinions on e.g. language innovation and literary genres, by analyzing relevant aspects of paratexts (prefaces, titles pages etc.), prescriptive language handbooks, genre prescriptions, archives (and consulting the various consisting studies on these topics, Sijs&Willemijns 2009; Sijs 2004; Wal 1995).

Historical sociolinguists usually define registers by general and broad sociological labels, such as gender. They, for example, distinguish registers used to address male versus female audiences. This subproject expands this usual historical sociolinguistic approach by adding text analytic approaches and concepts from literary studies to investigate how literary strategies and conventions interact with linguistic variation in different registers. By involving literary and textual factors, we will gain a more complete insight into the variation between and within different registers. For example: while Coornhert has been identified as an author who generally used post-nominal genitives (where the genitive is placed behind the substantive) (Damsteegt 1978), it has gone unnoticed that he also used pre-nominal genitives. Our preliminary research reveals that he opened his *Zedekunst* (1586) with various genitives related to God ('Godes ghoedheyd' [God's goodness] versus 'de Vader des lichts' [Father of light]). It is possible that the metaphorical character of the text influenced the way Coornhert used the variation available to him: since 'Father of light' denotes God on a figurative level, Coornhert used post-nominal genitives in the case of a metaphor.

The combination of approaches results in the following analytic model:



Corpus

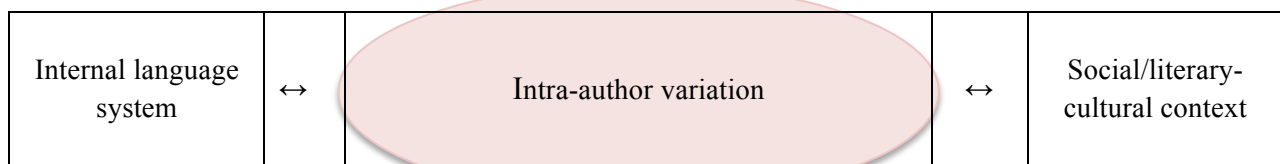
We will select case studies (authors, texts) that cover all of the axes described in I.2, with special attention for authors who consciously used and extended the variation options available in their language. We carefully select sets of texts that allow for comparison, for a better interpretation of analytic results: e.g. we compare Hooft's personal, handwritten letters with his printed, literary letters (Marion 2005) to investigate the effect of media on intra-author variation, and his earlier prose works to his later *Nederlandsche Historien* ('Dutch Histories', Hooft 2007) to investigate whether growing proficiency had any effects on intra-author variation. We will also include rhymed and metrical texts (like poetry, songs and plays) to get an even more informed view of textual variation in various genres and linguistic contexts.

Subproject 3: Quantifying intra- and inter-author variation in the Dutch Golden Age

(744 words)

Questions, Goals

This subproject pursues a quantitative approach to test the analyses developed in subprojects 1+2, as well as to enhance them by detecting additional variation patterns. It discloses variation patterns (i.e. the typology) as well as the frequency and distribution of these patterns within and between authors, registers and genres. To recall Hooft's use of negative concord again: subproject 3 reveals for instance how Hooft's emphatic use of negative concord is distributed with respect to text type, how frequent Hooft's use of a certain type of negation is in various genres, and how Hooft's language compares to, for instance, Vondel's.



In order to perform this quantitative research, existing digital tools for the enriching (e.g. tagging, parsing and searching) of texts have to be adjusted to 17th-century Dutch. Therefore, this project combines infrastructural and empirical subquestions:

Research Question
Which patterns, frequencies and distribution of (morpho)syntactic variation within and between authors and genres can be uncovered using digital tools?

Subquestions
A. Methodological subquestion How can we enrich 17th-century Dutch texts in such a way that they can be used to digitally analyze morpho-syntactic variation?
B. Empirical subquestions 1. What are the variation patterns within texts written by the same author? 2. [...] within the different genres used by the same author? 3. [...] between authors/genres?

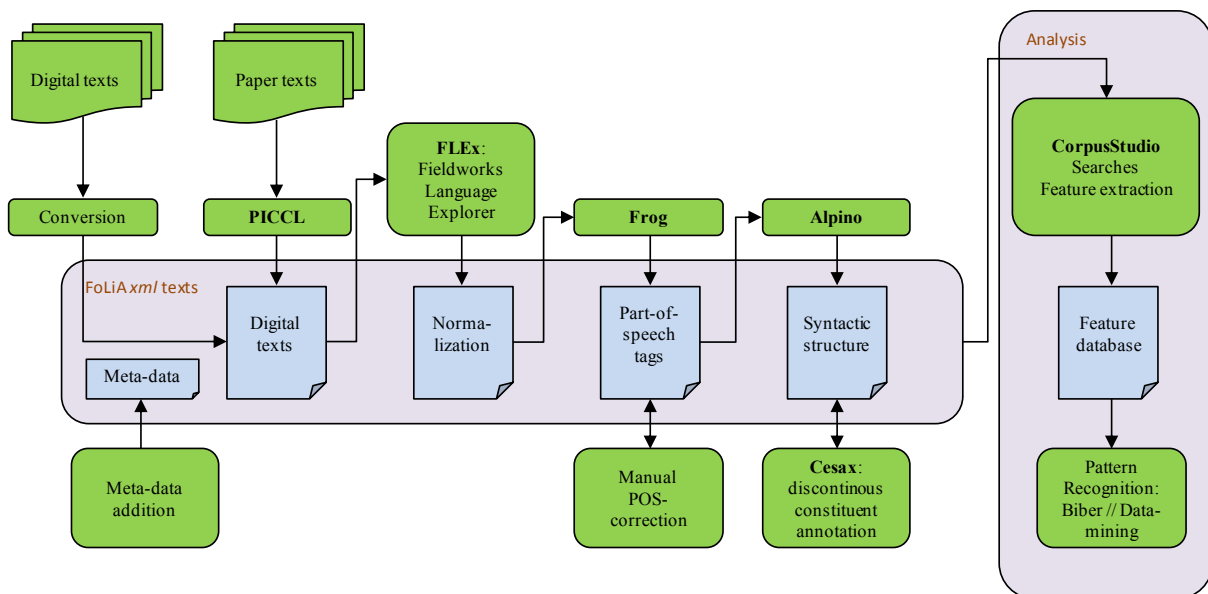
Ad Methodological subquestion:

To answer the methodological subquestion, we:

1. develop a syntactically parsed 500k-word text corpus of 17th-century Dutch texts;
2. enrich this text corpus with (morpho)syntactic information.

The corpus (primarily composed of already digitally available texts) and tools to enrich and search this corpus are developed in cooperation with [Nederlab](#) and [CLARIAH](#). The data will be made [CLARIN-compatible](#) and hosted at the Meertens Institute, a CLARIN-center, to ensure their sustainability.

We will pursue the following methodology to answer the methodological research question. Texts only available as scans will be OCR-ed (Optical Character Recognition) with the application PICCL (Reynaert 2014), which includes automatic error correction. These results are semi-automatically enhanced with the applications Fieldworks or Toolbox (SIL-International 2012), that add, as Hupkes (2014) and Koolen et al. (2006) propose (and also pursued in Nederlab and [CREATE](#)), a normalized layer of ‘modern’ Dutch to the 17th-century Dutch texts. This normalized layer will then be further enriched with part-of-speech tags, lemmas and a morphological analysis with the application Frog (Bosch et al. 2007). The part-of-speech tags will be checked and corrected semi-automatically, based on the accuracy levels provided by Frog. The Alpino-parser, next, adds syntactic annotation, arriving at a broad constituency level. The Cesax application (Komen 2011), as a final step, enhances the syntactic annotation by adding information about discontinuous constituents. We will estimate the accuracy of the resulting syntactic annotation and determine whether an additional correction step is needed depending on the results. The resulting texts in the FoLiA format (Gompel&Reynaert 2013) and added metadata can be searched in detail with the application CorpusStudio (Komen 2009), that has the additional benefit of enabling the inclusion of crucial text-external metadata (such as information on the author, year, genre) in the querying texts. This whole process enables us to measure characteristics of texts at the word level (e.g. the number of adpositions) as well as at the level of syntax (e.g. the number and type of constituents preceding the finite verb).



Ad Empirical subquestions:

We use the enriched text corpus to

1. search for (morpho)syntactic patterns, frequencies and distributions;
2. relate these patterns, frequencies and distributions to author and genre.

In addition to the (morpho)syntactic features that the other two projects investigate, this project will also include the (morpho)syntactic properties that have been uncovered in previous quantitative work on author/register variation by Biber (1995) and Fricke (2007), like the person/number/gender of pronouns, tense, aspect, and type of adverbial.

The project will determine which subsets of features characterize (a) an author, (b) a genre or (c) a combination of these factors. As such this subproject quantitatively tests the analyses developed in subprojects 1 and 2, and in turn feeds these projects by suggesting additional variation patterns to be analyzed in the other two subprojects.

15. Summary for the Public

(50 words)

In de Gouden Eeuw was het Nederlands volop in beweging. Daardoor was het taalgebruik opvallend gevarieerd: Nederlanders gebruikten ouderwetse en nieuwe taalconstructies probleemloos door elkaar. Hoe ontstonden die mogelijkheden in de taal van zeventiende-eeuwers? Hoe werden ze benut door taalkunstenaars als P.C. Hooft of in reisverslagen van Michiel de Ruyter?

16. Summary for non-specialists and the public

(800 words)

Taaldynamiek in de Nederlandse Gouden Eeuw

De Nederlandse Gouden Eeuw was een dynamisch tijdperk waarin er vernieuwingen plaatsvonden op vele terreinen, zoals cultuur, religie, wetenschap en handel. Ook de taal was volop in beweging. Het Nederlands, de eenheidstaal van de nieuwe Republiek, werd in steeds meer domeinen van de samenleving gebruikt (zoals het religieuze en wetenschappelijke domein) en er werden vele pogingen ondernomen om de positie van de moedertaal te versterken en de taal te standaardiseren. Zo kwamen medewerkers aan de Statenvertaling, afkomstig uit verschillende regio's, formele taalafspraken overeen. Ook natuurlijke taalontwikkelingen hadden een impact op het Nederlands: steeds meer eigenschappen uit het Middelnederlands (bijvoorbeeld naamval) verdwenen om plaats te maken voor nieuwe eigenschappen (zoals het gebruik van voorzetselgroepen).

Deze taalontwikkelingen resulteerden in veel variatie binnen het taalgebruik van auteurs. Het taalsysteem van een zeventiende-eeuwer bevatte bijvoorbeeld drie manieren om uit te drukken dat vader een broek bezit: *vaders broek*, *de broek des vaders* en *de broek van vader*. Dit project gaat in op de vraag waarom in een bepaalde situatie een van deze opties werd gebruikt. Hing dat samen met de regels van het genre, de afspraken die werden gemaakt in het Statenbijbelproject of de sociale achtergrond van het

beoogde publiek?

Dit project onderzoekt deze nog zo weinig bestudeerde *intra-author variation*, en wil begrijpen welke factoren die variatie tot stand brachten: hoe ontstond de *intra-author variation* in het zeventiende-eeuws Nederlands? Onze hypothese is dat de variatie het resultaat was van een dynamische interactie tussen het *interne taalsysteem* van taalgebruikers enerzijds en hun *sociaal/literair-culturele context* anderzijds. Het taalsysteem van een taalgebruiker maakte variatiemogelijkheden beschikbaar, die vervolgens door een taalgebruiker systematisch en vaak strategisch werden ingezet, afhankelijk van bijvoorbeeld het publiek of de doelstellingen en literaire vormgeving van zijn tekst.

We zullen een voorbeeld geven van het type verschijnselen waar we naar kijken, namelijk negatie. In het Middelnederlands werden zinnen ontkennend gemaakt door tweeledige negaties van het type *en...niet* (vergelijk het Franse *ne...pas*). In de zeventiende eeuw maakte deze vorm van negatie langzaam plaats voor eenledige negaties: *ic en sal niet moghen gaen* werd steeds vaker *ik zal niet mogen gaan*. Hooft gebruikte in zijn brieven eenledige en tweeledige negaties door elkaar. Was dat toeval of zit er systematiek achter zijn keuzes? Op basis van ons vooronderzoek denken wij dat in het interne taalsysteem van Hooft de twee typen negaties een verschillende lading kregen. Zo gaf Hooft zinnen van het type ‘niet dit, maar dat’ extra nadruk door de tweeledige ontkenning toe te voegen: ‘*Ick en zoek de rouw niet, maer zij weet mij te vinden*’ (Hooft 1624). Daarnaast lijkt de sociaal-culturele context relevant te zijn voor zijn keuzes. Juist in zijn brief aan zijn geliefde Eleonora Hellemans, waarin hij haar afwijzing betreurt, komen dubbele negaties vaker voor dan in andere brieven.

Hooft bezit dus blijkbaar een ‘subgrammatica’ – ofwel een ‘register’ – met twee negatie-varianten dat hij inzet in bepaalde situaties. In andere gevallen gebruikt hij juist zijn register met alleen eenledige negaties. Hoe dat werkt, willen we in dit project onderzoeken via drie samenhangende deelprojecten, waarin teksten van verschillende auteurs worden onderzocht. We kijken bijvoorbeeld zowel naar zowel taalkunstenaars die het Nederlands actief wilden vernieuwen en verfraaien (zoals Bredero en Hooft) als naar taalgebruikers die geschreven taal inzetten als een praktisch communicatiemiddel (Michiel de Ruyter), en combineren mannen met vrouwen, en migranten met in de Republiek geboren en getogen auteurs.

De eerste twee deelprojecten, gericht op de kwalitatieve analyse van *case studies*, worden uitgevoerd door AiO’s. Deelproject 1 verklaart *intra-author variation* vanuit het interne taalsysteem en deelproject 2 vanuit de literair-culturele context. Het postdoc-project legt op grootschalige wijze patronen van *intra-author variation* bloot, om zo de resultaten van de andere deelprojecten te testen en nieuwe variatiepatronen op het spoor te komen. Om dit kwantitatieve onderzoek uit te voeren, zal de postdoc nieuwe *tools* ontwikkelen om teksten uit het zeventiende-eeuws Nederlands te voorzien van syntactische informatie en automatisch te doorzoeken op syntactische structuren. De resultaten van deze drie deelprojecten zullen worden samengebracht in een aantal synthetiserende studies.

Innovatief aan dit project is de interdisciplinaire aanpak: het combineert methodiek en theorievorming uit de theoretische taalkunde, historische sociolinguïstiek, computationele taalkunde en vroegmoderne letterkunde. De variatie binnen taalgebruikers wordt normaal gesproken door theoretisch taalkundigen verklaard vanuit het taalsysteem en binnen de historische sociolinguïstiek vanuit sociale variabelen zoals geslacht en sociale klasse. Dit project brengt de beide perspectieven samen, en voegt daar een letterkundige benadering aan toe, om op die manier beter inzicht te krijgen in de literaire aspecten van taalvariatie binnen auteurs (zoals genreconventies) en in het strategische en creatieve gebruik van taalvariatie. Daarnaast profiteert dit project van recente ontwikkelingen in de computationele taalkunde. Dankzij de vernieuwende interdisciplinaire aanpak werpt dit project licht op zowel de grammaticale

kenmerken als de literaire en culturele factoren die taalvariatie tot stand brachten in een tijd van intensieve taalontwikkeling.

Literature (2p. A4 in 9pt)

- Adger, D., 'Variability and Grammatical Architecture'. In: C. Piccolo, ed., *Linguistic Variation in the Minimalist Framework*. Oxford, 2014.
- Adger, D. & J. Smith. 'Variation and the Minimalist Program'. In: L. Cornips & K. Corrigan, eds., *Syntax and Variation: Reconciling the Biological and the Social*, 149-178, Amsterdam, 2005.
- Adger, D. & J. Smith, 'Variation in agreement: A lexical feature-based approach', *Lingua* 120 (2010), 1109-1134.
- Auer, A. 'Precept and Practice: The Influence of Prescriptivism on the English Subjunctive'. In C. Dalton-Puffer et al., eds., *Syntax, Style and Grammatical Norms: English from 1500-2000*. Linguistic Insights. Frankfurt: Peter Lang, 2006, 33-53.
- Auer, A. 'Stylistic Variation'. In: A. Auer et al, ed., *Letter Writing and Language Change*. Cambridge, 2015, 133-155.
- Barbiers, S., 'Locus and limits of syntactic microvariation', *Lingua* 119-11 (2009), 1607-1623.
- Barbiers, S., 'Syntactic doubling and deletion as a source of variation'. In: C. Piccolo, ed., *Linguistic Variation in the Minimalist Framework*. 1e ed, Oxford 2014, 197-223.
- Barend M. & A.J. Gelderblom, eds., *Buyten gaets. Twee burlieke reisbrieven van Aernout van Overbeke*. Hilversum 1998.
- Belder, M. de & M. van Koppen, 'What may cause idiolectal variation: the case of Dutch ANN-compounds'. Paper presented at WCCFL 32, March 7-9 2014, Los Angeles, USC.
- Biber, D., *Dimensions of register variation: A cross-linguistic comparison*. Cambridge, 1995.
- Biberauer, T., 'Introduction'. In: T. Biberauer, ed., *The limits of syntactic variation*. Amsterdam, 2008, 7-72.
- Biberauer, T. & M. Richards, 'True optionality: When the grammar doesn't mind'. In: C. Boeckx, ed., *Minimalist Essays*. Amsterdam, 2006, 35-66.
- Borer, H., *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht, 1984.
- Borer, H., *In Name Only. Structuring Sense*. Volume I, Oxford, 2005.
- Bosch, A. van den et al, 'An efficient memory-based morphosyntactic tagger and parser for Dutch'. In: F. van Eynde et al, eds., *Selected Papers of the 17th Computational Linguistics in the Netherlands Meeting*. Leuven, 2007, 99-114.
- Branden, L. van den, *Het streven naar verheerlijking, zuivering en opbouw van het Nederlands in de 16de eeuw*. Gent 1956.
- Chomsky, N., *The Minimalist Program*. (Current Studies in Linguistics 28) Cambridge, 1995.
- Chomsky, N., 'Minimalist Inquiries: The Framework'. In: R. Martin et al., eds., *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*. Cambridge, 2000, 89-155.
- Damsteegt, B.C., 't'Samenvoeging in de Hertspiegel', *Tijdschrift voor Nederlandse Taal- en Letterkunde* 94 (1978).
- Damsteegt, B.C., *Van Spiegel tot Leeuwenhoek. Syntaktische en stilistische verschijnselen in 17^e-eeuwse teksten*. Leiden, 1981.
- Frieke-Kappers, Claertje, *The Creative Use of Genre Features. Continuity and Change Patterns of Language Use in Budu, a Bantu Language of Congo (Kinshasa)*. PhD dissertation, Free University, Amsterdam.
- Frijhoff, W.Th.M. & M.Spies, *Dutch Culture in a European Perspective 1650, hard-worn unity*. The Hague, 1999.
- Geerdink, N., *Dichters en verdiensten. De sociale verankering van het dichterschap van Jan Vos (1610-1667)*. Hilversum, 2012.
- Gelderen, E. van, *Cyclic Change*. John Benjamins, 2009.
- Gelderblom, A.J., "'Nieuwe stof in Neerlandsch". Een karakteristiek van Coornherts proza'. In: H. Bongers et al., eds., *Dirck Volckertszoon Coornhert. Dwars maar recht*. Zutphen, 1989, 98-114.
- Gianollo, C. et al, 'Three fundamental issues in parametric linguistics', *The limits of syntactic variation* (2001), 109-142.
- Gompel, M. van & M. Reynaert, 'FoLiA: A practical XML Format for Linguistic Annotation-a descriptive and comparative study', *Computational Linguistics in the Netherlands Journal* 3 (2013), 63-81.
- Goss, E. *Negotiated language change in Early Modern Holland. Immigration and linguistic variation in The Hague (1600-1670)*. PhD-dissertation, University of Wisconsin, Madison 2002.

- Goss E. & R. Howell, 'Social and structural factors in the development of Dutch urban dialects in the Early Modern period'. In: T. Cravens, ed., *Variation and reconstruction*. John Benjamins 2006, 59-88.
- Greenblatt, S., *Renaissance Self-Fashioning: from More to Shakespeare*. Chicago, 1980.
- Hendriks, J., 'Case marking (accounts) in collapse: Evidence from Early Modern Dutch egodocuments (1572-1573)'. In: M. Ponsonnet, L. Dao & M. Bowler (eds.), *Proceedings of the 42nd Australian Linguistic Society Conference – 2011, Australian National University, Canberra ACT, 2-4 December 2011*, 2012, 123-151.
- Hoeksema, J., 'Negation and negative concord in Middle Dutch'. In: D. Forget et al., eds., *Negation and Polarity: Syntax and Semantics*. Amsterdam, 1987, 139-158.
- Hooft, P.C. *Nederlandse Historiën. Een keuze uit het grote verhaal van de Nederlandse Opstand*. Ed. F. van Gestel et al. Amsterdam 2007.
- Hooft, P.C., brief aan Hellemans, 27-08-1627. In: H.W. van Tricht, ed., *De briefwisseling van P.C. Hooft*. Culemborg, 1977.
- Horst, J. van der & M. van der Wall, 'Negatieverschijnselen en woordvolgorde in de geschiedenis van het Nederlands'. In: *TNTL* 95-1 (1979), 6-37.
- Horst, J. van der, *Geschiedenis van de Nederlandse syntaxis*. Leuven, 2008.
- Hulshof, Casper, 'Leren schrijven met de hand is essentieel. De kracht van "embodied cognition"'. Blogbijdrage, 11 maart 2013. <http://onderzoekonderwijs.net/2013/03/11/leren-schrijven-met-de-hand-is-essentieel-de-kracht-van-embodied-cognition/>
- Hupkes, D., 'Part-of-Speech Tagging van asynchrone Bijbelteksten'. 10 oktober 2014, intern onderzoeksrapport.
- Jansen, J., *Imitatio. Literaire navolging in de Europese letterkunde van de renaissance*. Hilversum, 2008.
- Jansen, J., 'Het proza'. In: In: G.A. Bredero, *Proza*. Ed. Jeroen Jansen. Hilversum 2011, p. 34-66
- Jespersen, O., *Negation in English and other languages*. Copenhagen, 1917.
- Kayne, R., 'Some notes on comparative syntax with special reference to English and French'. In: G. Cinque & R. Kayne, *The oxford handbook of comparative syntax*. Oxford, 2005, 3-69.
- Kemenade, A. van, *Syntactic case and morphological case in the history of English*. Dordrecht: Foris, 1987.
- Kemenade, A. van, 'Jespersen's cycle revisited: formal properties of grammaticalization'. In: S. Pintzuk et al., eds., *Diachronic Syntax: Models and Mechanisms*. Oxford, 2000, 51-75.
- Koelmans, L., *Het Nederlands van Michiel de Ruyter: morfologie, woordvorming en syntaxis*. Assen, 2001.
- Komen, E., *CorpusStudio*. Nijmegen, 2009.
- Komen, E., *Cesax: coreference editor for syntactically annotated XML corpora*. Nijmegen 2011.
- Komen, E., *Finding Focus. A study of the historical development of focus in English*. PhD dissertation, RU Nijmegen, 2013.
- Koolen, M. et al., 'A cross-language approach to historic document retrieval', *Advances in Information Retrieval* (2006), 407-419.
- Koppenol, J., T. van Strien, m.m.v. N. Veldhorst, eds., *P.C. Hooft. De gedichten*. Amsterdam, 2012.
- Kroch, A., 'Reflexes of grammar in patterns of language change', *Language Variation and Change* 1 (1989), 199-244.
- Kroch, A., *Syntactic Change, in The Handbook of Contemporary Syntactic Theory*. Oxford 2001.
- Lefebvre, M., 'Voer je eigen taalonderzoek!'. In: A. Mottart & S. Vanhooren, eds., *28^{ste} HSN-conferentie*. Gent 2014, 184-188.
- Maas, W., 'Geen slappe thee, maar espresso!'. In: *Literatuurgeschiedenis op de havo. Themanummer Levende talen*, 2008, 7-9.
- MacKenzie, L., 'Intraspeaker change in later life: Implications for grammatical representations'. Paper presented at FWAV, University of Iceland, May 28, 2015.
- Marion, O. van, *Heldinnenbrieven. Ovidius' Heroides in Nederland*, Nijmegen, 2005.
- Mesthrie, R., J. Swann, A. Deumert & W.L. Leap, *Introducing Sociolinguistics*. Second edition. Edinburgh, 2013.
- Nevalainen, T. & H. Rumoulin-Brunberg, *Historical sociolinguistics: Language Change in Tudor & Stuart England*. 2003.
- Nevalainen, T. & H. Rumoulin-Brunberg, 'Historical sociolinguistics: Origins, motivations and paradigms.' In: J. Hernández-Campoy & J. Conde-Silvestre (eds.) *The handbook of historical sociolinguistics*. 2012, 22-40
- Neeleman, A. & F. Weerman. *Flexible syntax: A theory of case and arguments*. Dordrecht: Kluwer, 1999.
- Nobels, J. (Extra)Ordinary Letters. A View from below on 17th-century Dutch. PhD dissertation, Leiden, 2013.
- Overdiep, G.S., *Verzamelde opstellen over taal- en letterkunde*. Antwerpen, 1947-1948.
- Paardekooper, P., 'Bloei en ondergang van onbeperkt ne/en, vooral dat bij niet-woorden', *Neerlandistiek.nl* 2006.
- Quené, H., 'Longitudinal trends in speech tempo: The case of Queen Beatrix', *Journal of the Acoustical Society of America* 133 (2013), 452-457.
- Reynaert, M., 'Synergy of Nederlab and@ PhilosTEI: diachronic and multilingual Text-Induced Corpus Clean-up', *Proceedings of the 6th International Language Resources and Evaluation (LREC'14)*. Iceland, 2014
- Robben, P., 'Wetenschappelijke geschiedenis van de syntaxis spreekt leerlingen aan'. In: *Levende Talen Magazine* 97 (2010).
- Roeper, T., 'Strict interfaces and three kinds of multiple grammar'. In: E. Rinke & T. Kupisch, eds., *The Development of Grammar. Language acquisition and diachronic change*. In honour of Jürgen M. Meisel. 2011, 205-228.
- Schenkeveld-van der Dussen, M.A., *Dichter en boer. Hubert Korneliszoon Poot, zijn leven, zijn gedichten*. Amsterdam, 2009.
- Schilling-Estes, N., 'Investigating Stylistic Variation'. In: J.K. Chambers et al, eds., *Handbook of language variation and change*. Oxford 2002, 375-401.
- Sijs, N. van der, *Taal als mensenwerk. Het ontstaan van het ABN*. Den Haag, 2004.
- Sijs, N. van der & R. Willemijns, *Het verhaal van het Nederlands*. Amsterdam, 2009.
- SIL-International, 2012. *SIL FieldWorkds*: SIL International.
- Spies, M., *Rhetoric, rhetoricians and poets. Studies in renaissance poetry and poetics*. Amsterdam, 1999.
- Spitzer, M., *Digitale dementie. Hoe wij ons verstand kapotmaken*. Amsterdam, 2013.

- Thráinsson, H., 'Ideal Speakers and Other Speakers: The case of dative and some other cases'. In: B. Fernandez & R. Etxepare, eds., *Variation in datives: a microcomparative perspective*. Oxford, 2012, 161-188
- Tieken-Boon van Ostade I.M., 'Dr. Johnson and the auxiliary *do*''', *Folia Linguistica Historica* X/1 (1991), 145-162.
- Tieken-Boon van Ostade I.M., 'Flat adverbs and Jane Austen's letters'. In: M. van der Wal & G. Rutten, 2013, 91-106.
- Tricht, W.H. van, 'Iets over Leonora Hellemans' schijndomicilie te Zevenbergen', *TNTL* 73 (1955), 90-92.
- Wal, M. van der, 'Taalidealen, taalnormen en taalverandering: Johan de Brune in linguistisch perspectief'. In: P.J. Verkruijsse, *Johan de Brune de Oude (1588-1658): een Zeeuws literator en staatsman uit de zeventiende eeuw*. Middelburg, 1990, 54-68.
- Wal, M. van der, *De moedertaal centraal. Standaardisatie-aspecten in de Nederlanden omstreeks 1650*. Den Haag, 1995.
- Wal, M. van der & C. van Bree, *Geschiedenis van het Nederlands*. Utrecht, 2008.
- Wal, M. van der & G. Rutten, *Touching the Past. Studies in the historical sociolinguistics of ego-documents*. Amsterdam 2013.
- Wal, M. van der & G. Rutten, *Letters as Loot. A sociolinguistic approach to 17th- and eighteenth-century Dutch advances. Historical Sociolinguistics 2*. Amsterdam, 2014.
- Weerman, F., M. Olson & R. Cloutier, 'Synchronic variation and loss of case: Formal and informal language in a Dutch corpus of 17th-century Amsterdam texts'. In: *Diachronica* 30:3 (2013), 353-381.
- Witstein, S.F., *Funeraire poëzie in de Nederlandse Renaissance. Enkele funeraire gedichten van Heinsius, Hooft, Huygens en Vondel gezien tegen de achtergrond van de theorie betreffende het genre*. Assen, 1969.
- Wouden, T. van der, *Negative Contexts*. Doctoral dissertation, University of Groningen, 1994.
- Wouden, T. van der, 'Meer over dubbele ontkenningen. Reactie op Piet Paardekooper'. *Neerlandistiek.nl*, 2007.
- Yang, C., 'Internal and external forces in language change', *Language Variation and Change* 12 (2000), 231-250.
- Yang, C., 'Three factors in language variation', *Lingua* 120-5 (2010), 1160-1177.
- Zeijlstra, H., *Sentential negation and negative concord*. PhD dissertation, Utrecht, 2004.