LANGUAGE-INTERNAL AND -EXTERNAL FACTORS CONDITIONING A GRAMMATICAL CHANGE IN THE EARLY MODERN DUTCH VERB *ZIJN* 'BE' PARADIGM[*]

Christina Ilse Terpstra
University of Calgary

1. Introduction

Like other Germanic languages, Dutch exhibits suppletion in the verb root ZIJN 'be'. So-called *b*-roots (Proto Indo-European **bheu*) and *s*-roots (Proto Indo-European **h1es*) merged into a single paradigm (Donaldson 1983: 182). In Middle Dutch, (MiD), there were two alternating forms for the present indicative plural: *b*-root form *bennen* and *s*-root form *zijn*[1] (Donaldson 1983: 182). Speakers of most Modern Dutch (MoD) varieties, however, use the form *zijn* for all present indicative plural expressions, suggesting that *bennen* has become obsolete in MoD:

Table 1a. ZIJN paradigm in MiD

	sg.	pl.
1	*bem*/*ben*	***sijn*/*benne***
2	*best*/*bist*	***sijt*/*benne***
3	*es*/*is*	***sijn*/*benne***

(Donaldson 1983: 182)

Table 1b. ZIJN paradigm in MoD

	sg.	pl.
1	*ben*	***zijn***
2	*bent*	***zijn***
3	*is*	***zijn***

(Donaldson 1981: 139)

Where forms become uniform through analogy, the variants with greater usage provide the directionality for competing forms. Relevant to the case at hand, it is expected that paradigmatic uniformity spreads from the third person to the first and/or second person, because the third person is used with greater frequency in discourse (Kuryłowicz 1947, cf. Hock 1986: 183; Aalberse and Don 2009: 330, 2011: 346). However, a strictly language-internal analysis may not holistically account for a change's development (cf. Labov 1966; Weinreich et al. 1968). Recent studies in the domain of historical sociolinguistics have illustrated how social patterning influences historical developments. Social factors such as gender, age, and class (e.g. Blaxter 2015; Rutten and van der Wal 2014: 137-141; Blas Arroyo 2016: 22-23, respectively) have been identified as influencing the patterning of variation and linguistic change.

With the use of the *Brieven als Buit* ('Letters as Loot') corpus, I explore whether social (gender, age, class) and linguistic factors (subject-verb agreement) had an impact on

[*] I would like to acknowledge Drs. A. Pounder, D. Storoshenko, and L. Brown for their suggestions in improving the research and writing component of this project. I would also like to thank Dr. T. Fung for his assistance with the statistics.

[1] See section 3.2 on spelling variation of the *bennen* and *zijn* wordforms

Actes du congrès annuel de l'Association canadienne de linguistique 2018.
Proceedings of the 2018 annual conference of the Canadian Linguistic Association.

the developments towards a uniform plural ZIJN paradigm. The selected period of study is the Early Modern Dutch period (ca. 17th to 19th century; henceforth "EMD"), which is the the intermediate period following the attested presence of the *bennen*/*zijn* alternation (i.e., MiD) and preceding the period of the eventual obsolescence of *bennen* in (i.e., MoD).

The rationale of my investigation is that analogy only partly accounts for the obsolescence of *bennen*, and that it was further socially conditioned. The results of my analysis confirms this hypothesis and demonstrates that subject-verb agreement influenced the directionality of the shift in the paradigm, while social factors accounted for the actuation of the shift. The results suggest that the obsolescence of *bennen* is most presumably due to it being a colloquial variant, and as a colloquial variant, its usage was disfavoured to the prestigious *zijn* form. My investigation illustrates the importance of going beyond published texts to bring certain insights to light; specifically, this study is unparalleled in that it is the first to provide a thorough empirical analysis examining the linguistic and social factors involved in the loss of *bennen*, a variant with few mentions in the literature. More broadly, it contributes to the growing understanding that language change must be examined both from a language–internal and –external perspective.

The rest of the paper is structured as follows: section 2 serves as the background for this investigation, section 3 provides the methodology and hypotheses, section 4 lays out the results of the analysis which are subsequently discussed in section 5, and section 6 offers concluding remarks.

2. Background

As mentioned in section 1, the verb ZIJN 'be' is suppletive, and its paradigm contains alternating *b*-roots and *s*-roots. Van den Toorn et al. (1997: 58, 118) provide the Old Low Franconian (Old Dutch, henceforth OLF) and Middle Dutch paradigms:

Table 2a. ZIJN paradigm in OLF

	sg.	pl.
1	*bim*/*bin*	***siin***
2	*bis*/*bist*	***siit, sint***
3	*is*/*ist*	***sint***

van den Toorn et al. (1997: 58, 118)

Table 2b. ZIJN paradigm in MoD

	sg.	pl.
1	*bem*/*ben*/*bin*	***sijn***
2	*bes*(*t*)/*bis*(*t*)	***sijt***
3	*es*/*is*	***sijn***

van den Toorn et al. (1997: 58, 188)

Van den Toorn et al. (1997: 58) report that *s*-root forms were used for plural expression as early as the 10th century. As is clear from Table 2a-b, only minor developments took place in between Old Low Franconian and Middle Dutch. At some point between Middle Dutch and Modern Dutch, the contemporary second plural pronoun *jullie* 'you (pl)' replaced *gij* in most Netherlandic dialects. Along with this replacement, the corresponding word-form *sijt* was replaced by *zijn* (Donaldson 1983: 182). As for the development of the *b*-root plural form *bennen*, very little information is available in the literature. It is unclear when *bennen* entered the paradigm. The first attestation of its presence is mentioned by

Donaldson to be in Middle Dutch. Donaldson describes the form *bennen* to belong to the "plat" varieties or colloquial non-standard varieties, and he states that it was "heard throughout the plural" (1983: 182). As to the later period, Early Modern Dutch, Koelmans (1978: 24) notes that in the sixteenth to seventeenth century, the colloquial *bennen* appeared for the first and third person, though he does not report on its use with the second person plural.

To my knowledge, no other works besides Donaldson (1983) and Koelmans (1978) mention *bennen*, nor does either author provide an empirical analysis of this alternation. According to Elspaß (2012: 161, 2007: 3), historical linguists have had the tendency of focusing their analysis on the dominating norm, by limiting their scope to published texts, ignoring linguistic input from minority languages, non-standard varieties, and registers. My assumption is that if the form *bennen* was indeed a colloquial variant, scholars may either not have encountered the form in their analyses of official texts or have intentionally dismissed it as irrelevant.

2.2 Language history "from below" approach in Historical Sociolinguistics

Analyzing historical developments touching on colloquial registers, like the *bennen* vs. *zijn* alternation addressed here, is a major goal of historical sociolinguistics. Historical sociolinguistics attempts to address the over-representation of dominant registers by employing the language history "from below" approach. This approach, examines linguistic patterns produced by the marginally–represented majority population which includes farmers, artisans, soldiers, and housemaids (Elspaß 2007: 5), in addition to the well represented upper-classes. The view "from below," Elspaß states, "implies a change of perspective from a 'bird's eye' view to a 'worm's eye' view" (2012: 160), by which is meant that language development is observed through utilizing media accessible to a wider authorship. By assuming a "worm's eye view" of language (Elspaß 2012: 160), we can identify if changes are "from above" the level of awareness, that is, speakers are aware of the variation because variants have differing prestige assignments, and or "from below" the level of awareness, that is, variants speakers are not aware of the variation, and variants do not have a social evaluation.

Furthermore, it is important to examine language variation and change beyond published texts because variation is structured differently for spoken and written registers (D'Arcy and Tagliamonte 2015). Since there are no recordings from before the 19th century, the proposed text types for analyzing Early Modern Dutch are so-called "ego-documents," constituting private letters, diaries, inventories, cookbooks, petition letters, and amateur theatrical schemes (Elspaß 2007: 5). Ego-documents, present one of the closest alternatives to the vernacular language in the written medium (Elspaß 2012: 158). They present relatively high degrees of orality, informality and unplannedness (cf. Koch and Oesterreicher 1994 on "texts of immediacy"), and they were not intended for a large audience; as such, there is no pressure to abide by any literary conventions, and therefore present a more vernacular register (Dossena 2007, Kielkiewicz-Janowiak 2012: 309).

2.3 Possible linguistic factors for the restructuring of the ZIJN paradigm

Analogy results morphologically, syntactically, or semantically related elements becoming more similar to each other (Hock 1986: 167). Kuryłowicz (1947) proposes that the basic form provides the core of the derived forms. "Basicness" is understood in terms of spheres of usage: items which are used in more contexts or with greater frequency are said to be more basic (Hock 1986: 183). For example, the third person is considered to be more basic than the first or second person (Hock 1986: 217-220). In parallel, more recent studies, such as those by Aalberse and Don (2007, 2009, 2011), refer to this phenomenon of analogy as "paradigmatic neutralization." Similar to Kuryłowicz, Aalberse and Don claim that the form used to express third person in regular verbs overtook the forms expressing other person marking because of the higher frequency of third person usage in speech (2009: 330, 2011: 346).

3. Methodology

3.1 Corpus

To identify the developments of the *bennen* vs. *zijn* alternation in Early Modern Dutch, the *Brieven als Buit* corpus was selected. This corpus consists of approximately a thousand private letters written between the 17th and 19th century. These letters served as correspondence between sailors and colonizers and their families during the Golden Age. These letters were confiscated by English pirates (hence the name *Brieven als Buit* or 'letters as loot'). The letters are organized into an electronic corpus with word token for lemma annotation, allowing for efficient data collection (see Rutten and van der Wal 2014). Additionally, the letters are organized by date and annotated for the author's age, gender, class, and place of origin, when this information is available. What makes this corpus an exciting source for historical linguistics is that it contains letters from individuals of all classes, along with an evenly distributed sample of letters from both men and women. The diversity in its senders allows for documenting the colloquial speech of the lower, middle, and upper classes. Furthermore, cross comparison between various social groups allows for socio-linguistic nuances to become apparent.

3.2 Data collection

In this study, I examine the *bennen* vs. *zijn* word-form alternation of the lexeme ZIJN 'be' in the 17th and 18th century. Data was collected by searching the corpus for instances of *bennen* and *zijn*. Each instance of these wordforms in the corpus counted as a token for which information of the independent variables, gender, age, and class, subject agreement, and time was recorded. The corpus contains enough tokens for both word-forms to perform a quantitative analysis. No additional morphological alternates for the *zijn*/*bennen* alternation were found. In the collecting of tokens, spelling variation was a cause of caution, as this variation might express phonological/morphological distinctions. However, spelling variation primarily concerned the orthographic means to express the vowel (e.g. <*b**e**nne*>, <*b**ei**nnen*>, <*b**i**nne*>, <*b**y**nne*> for *bennen,* <*z**ij**n*>, <*z**y**n*>, <*z**eij**n*>, and <*z**i**n*> for *zijn*) as well as the sibilant in the onset (e.g. <***s****ijn*> and <***z****ijn*>, for *zijn*), and had no

influence on the distribution of the dependent variables across linguistic factors. Tokens presented in parallel constructions, i.e. *zijn [x] and zijn [y]*, as opposed to elliptical constructions ,i.e. *zijn [x] and [y]*, were both only counted as one instance.

In order to quantitatively measure the significance of differences in the distribution of the dependent variable for the linguistic and social factors, the Generalized Estimated Equation (GEE) test was selected. The selected test is appropriate for this type of data organization because, first of all, the data are organized into a categorical/discrete format in the form of token counts in order to record the frequency of occurrence; GEE deals with the probability of occurrences. Secondly, GEE is equipped to cope with "repeated measures data." Varied quantities of tokens were provided by the speakers, i.e., speakers provided minimally one relevant token to maximally 27 tokens. The GEE is capable of factoring out the "noise" from intra-speaker variation (Zeger and Liang 1986: 13). Thirdly, unlike *t*-tests which need balanced comparison samples, the GEE procedure can compare samples of uneven sizes. Not all of the social factors were available for each speaker, thus a test was needed that would be able to cope with missing data. Results from the GEE analyses were interpreted in terms of measuring to what extent the data distribution deviates from the expectation, or null hypothesis with a threshold of $p = \geq 0.05$ α. The GEE analysis was performed using SPSS software.

3.3 Independent variables

Time: the dates of the letters were recorded to measure the difference in distribution of *zijn* and *bennen* over time. The dates of the letters are based on the dates provided by the authors themselves. Due to time constraints, token collection was restricted to two five-year periods: Period 1 (P1), encompassing letters written between 1664 and 1669, and Period 2 (P2), encompassing letters written between 1779 and 1783.

Subject-verb agreement: Like English, Dutch has a three-way person distinction, and a two-way number distinction. Although Donaldson (1983: 182) states that *zijn* and *bennen* pertain to plural expression, where applicable, I record usage of *bennen* and *zijn* with all the possible person-number subject agreement combinations. Thus, the variable values include first, second, and third person in the singular or plural.

As for social factors, I recorded information regarding the author's gender, age, and class, where this information was available. The independent variable, *Gender*, was added as an external factor, since numerous (historical) sociolinguistic studies (e.g. Rydén and Brorström, 1987: p. 200–207; Labov, 1990; Kytö, 1993; Nevalainen and Raumolin-Brunberg, 2003; Blaxter, 2015) have observed a gendered patterning in linguistic variation. The corpus developers state that the authors' genders were identified with relative ease, as they mostly inferenced the gender based on the authors' names (Rutten and van der Wal 2014: 9). Only one letter in my data collection did not have this information available (the letter was anonymous).

In terms of the variable *class*, my study follows the class stratification employed by the corpus, which is based on Frijhof and Spies' Republic of Seven United Provinces convention (Frijhof and Spies 1999: 188-190; cf. Rutten and van der Wal 2014: 9-10). The corpus developers use a four-way class system including the low class ("L": blue-collar

workers e.g. sailors, servants, or soldiers), low-middle class ("ML": petty bourgeoisie, e.g. small-shop keepers, small-scale craftsmen, or minor officials), high-middle class ("MH": large-store keepers, or non-commissioned officers), and a high class ("H": bourgeoisie e.g. wealthy merchants, ship owners, academics, or commissioned officers). While Frijhof and Spies make a six-way class distinction, the developers of the corpus collapsed these into four classes, omitting the extremely poor and the extremely rich; the former, which includes tramps, beggars, and the disabled did not normally have access to written communication and are therefore not represented, while the latter, is omitted because of over-representation in external corpora (Rutten and van der Wal, 2014, p. 10). Lastly, female authors were assigned a class label according to their husband's/father's class division.

Lastly, *Age* was included as a social factor. Studies identifying age-conditioned linguistic changes are relatively rare. However, the few studies that are available fall into two opposing categories: a) studies reporting age-patterning which parallels change over time (viz. apparent-time as opposed to real time change) (e.g. Labov 1972; Chambers 2002), and b) those which identify a difference between youth-speak and adult-speak (cf. Kerswill 1996, Blas Arroyo 2016: 22-23). For the current investigation, I utilize the age-categorization provided by the corpus. The corpus developers organized *Age* into three cohorts: thirty years or below, between thirty and fifty years, and fifty or above. Since letters often do not provide straightforward information on age, a number of inferential clues were used by the corpus developers. For some of the letters, external documents such as marriage licenses, baptismal and burial registers were used to approximate the age of the author. Furthermore, "stage of life" clues were inferred from letters which specified kinship relationships between the author and the addressee/third party. For example, parent and grandparent status were grouped in the thirty-to-fifty and the fifty-or-above category respectively (Rutten and van der Wal 2014: 9).

The variable *Authorship* was added as a control factor; Dutch literacy in the context of 17th to 18th century lies on a continuum. At that time, reading was usually taught before writing and not simultaneously. Part of the population could neither read nor write. In order to communicate with their loved ones overseas, authors sometimes hired scribes to write the letters for them (Rutten and van der Wal 2014: 14). While these scribes were often aware of dialectal differences in writing, it is still a cause for caution when it comes to the representability of the text for colloquial tendencies. The corpus developers used a number of clues such as external documents, cross comparison between letters, and in-text clues to determine self vs. scribal authorship (see Rutten and van der Wal 2014). My preliminary statistical analysis of authorship effect, however, was shown to be inconclusive, meaning that I could safely proceed with including both self-authored and scribal-authored letters in my data sampling.

3.4 Hypotheses

Based on the observations provided by the relevant literature, the following hypotheses arise: the obsolescence of *bennen* should be evident from a spread of the third person (which is hypothesized to take on the form *zijn* since it is the form that eventually wins out) on to the first and second person because of its relative frequency (cf. Kuryłowicz

1947, see also Hock 1986: 167; Aalberse and Don 2009: 330, 2011: 346). In addition to this linguistic phenomenon, it is further predicted that social factors also play a role. Due to the fact that *bennen* has been described as a colloquial variant (Donaldson 1983: 182), it is predicted that it is socially stratified by the class index. It is likely that this variant was dis-preferred by the higher strata, and due to its discouraged use, it became obsolete. Secondly, as previous studies (Labov 1990; Blaxter 2015) have pointed out that women are usually the carriers of change, I predict that if it is patterned across gender, women are hypothesized to prefer the form *zijn* to a greater degree than men. Lastly, if the variants pattern by age, the variation either conforms to age-graded patterning (i.e. a youth-speak vs. adult-speak patterning, cf. Kerswill 1996; Blas Arroyo 2016: 22-23), or it reflects a generational change in apparent-time (vs. real-time, cf. Labov 1972; Chambers 2002).

4. Results and analysis

In total, 519 (71 for *bennen* and 448 for *zijn*) tokens for Period 1 and 516 (49 for *bennen* and 467 for *zijn*) tokens for Period 2 were identified. A total of 43 instances of *bennen* and *zijn* were identified for verbs with person marking in the singular, making up 4% of the data (43/1035). While their individual frequencies are too small to form part of the statistical analyses, the number of instances with singular person marking is nonetheless surprising. I identified a quarter of singular subject agreement cases (ten tokens) to occur with formulaic phrases. The phrase provided in (1) below, was used with minor variation for all ten instances:

(1) *een* *vrydelycke* *groetenijsse* ***sijn*** *gheschreven*
a friendly greeting are PRF-write
'a friendly greeting is written/ has been written (to you)'
(Magdalena Tijssen, September 1664: *Brieven als Buit* corpus)

As shown in (1), there is a number agreement mismatch between the singular subject noun *groetenijsse* and the supposedly plural verb form *sijn*. I suppose that the mismatch was due to it being a formulaic collocation, as such, it may have lost internal morphological transparency.

It was equally surprising that, in inverse, the corpus contained very few instances of the second person plural, only four to be exact. All four examples utilized the *zijn* variant, an example of which is provided in (2):

(2) *ik* *hoopen* *dat* *gij* *allen* *gezond* ***zyn***
I hope that you all healthy are
'I hope that you are all healthy'
(Isaac de Vijver, November 1780: *Brieven als Buit* corpus)

Since there are only four instances of the second person plural and the singular forms exhibit idiosyncratic behaviour, only the first person and third person plural were used as part of the statistical analyses on subject-agreement. For the first person plural, a total of 328 instances were found, 60 of which paired the first person plural agreement with *bennen* (18.3%), and 286 of which paired the first person plural agreement with *zijn* (81.7%). As

for the third person plural, a total of 643 instances were observed, 52 of which paired the third person plural agreement with *bennen* (8.1%), and 591 of which paired the third person plural agreement with *zijn* (91.9%). The statistical analysis shows that the distribution difference of *zijn* and *bennen* across the first person and third person plural is significant: $\chi^2(1)= 11.81$, $p < 0.01$. It is more likely for speakers to utilize the *zijn* alternate with the third person plural, than it is with the first person plural; the likelihood of selecting *zijn* with a first person referent is 0.83, while the likelihood to select *zijn* for the third person referent is 0.92. These results are provided in graphical format in Figure 1:

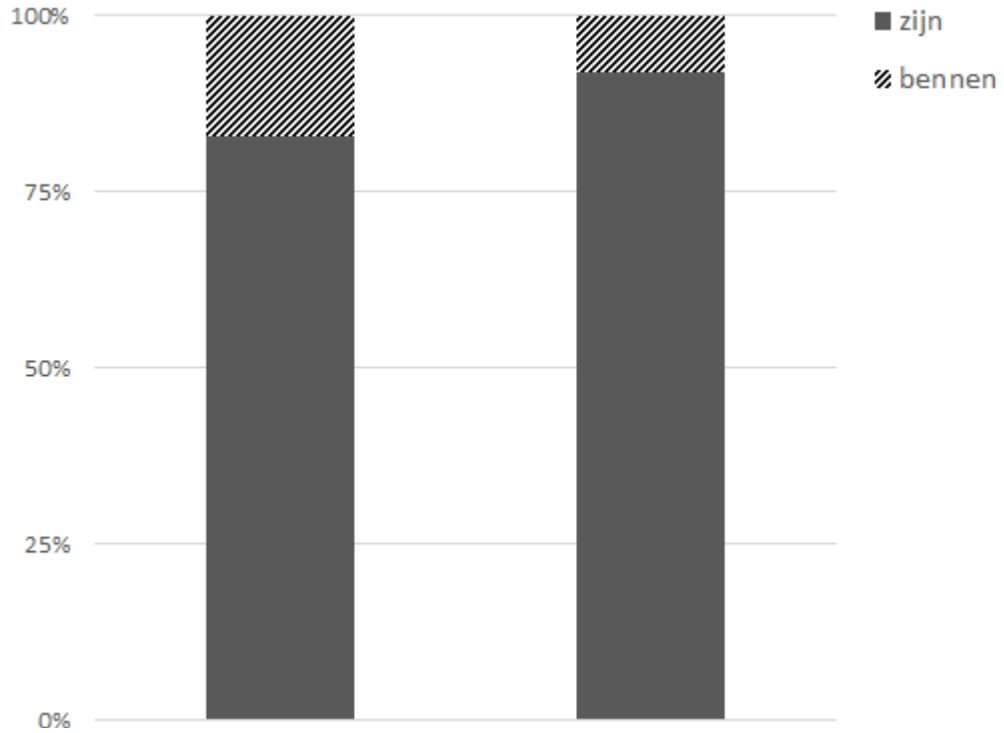

Figure 1. bennen *vs.* zijn *selection across first and third person plural*

The interaction between time and subject agreement was also considered. It was found that the distribution of *zijn* and *bennen* for subject selection did not differ significantly from Period 1 to Period 2: $\chi^2(1)= 2.09$, $p = 0.15$.

Turning now towards the analysis of social factors, a statistical analysis of gender determined that there is no significant distribution difference of *zijn* or *bennen* selection for men or women: $\chi^2(1)= 0.01$, $p = 0.94$. A statistical analysis of class, however, shows that *zijn* or *bennen* selection across the various social strata is significant: $\chi^2(3)= 14.05$, $p < 0.01$. The individual probability trends indicate that the likelihood for speakers to select *zijn* increases by class index: the likelihood of selecting *zijn* is 0.79 for the lower-class (L), 0.81 for the low-middle class (ML), 0.89 for the high-middle class (MH), and 0.96 for the higher-class (H). Figure 4 provides the full probability schema.

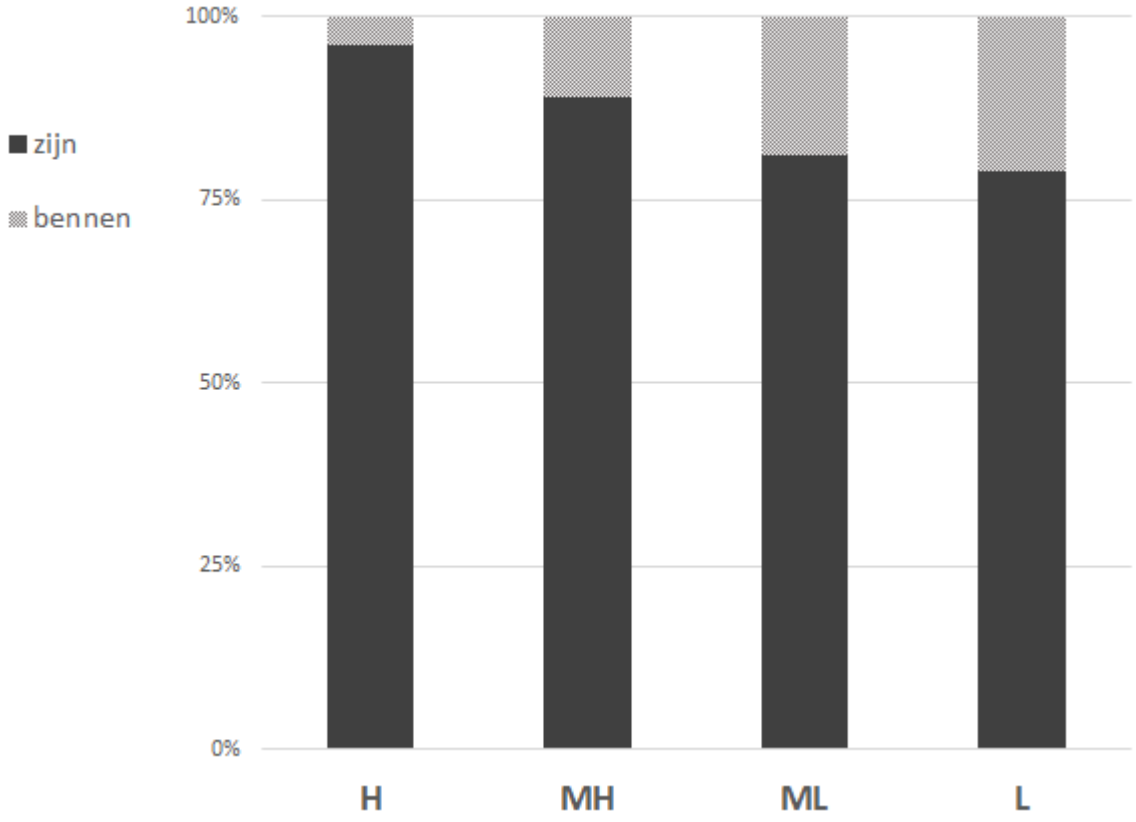

Figure 2. Selection of zijn *and* bennen *across social strata*

Results from an analysis of *zijn* vs. *bennen* selection for the three age cohorts (<30, 30-50, >50) indicate that the deviations in selection distribution per cohort are also significant: $\chi^2(2)= 7.03$, $p = 0.03$. Parallel to the results for class, the likelihood for speakers to select *zijn* increases by age index; the older speakers are, the higher the probability that they will select *zijn*. Speakers thirty years old and younger have a 0.83 probability of *zijn* selection, speakers between the ages of thirty and fifty have a 0.91 probability of *zijn* selection, and speakers fifty years and older have a 0.96 probability of *zijn* selection. Figure 3 provides the breakdown of *bennen* vs. *zijn* selection across the age cohorts.

In addition to main effects, I also performed analyses of the two-way interactions for the social variables. The interaction of gender and class does not indicate significant distribution differences in the selection of *zijn* or *bennen*: $\chi^2(3)= 3.13$, $p = 0.37$. The interaction of gender and age also does not indicate statistically significant distribution differences in the selection of *zijn* or *bennen*: $\chi^2(2)= 2.69$, $p = 0.26$. Unfortunately, the interaction of class and age could not be performed due to insufficient token counts for all of the class by age combination cells.

Lastly, I performed a general distribution comparison of *zijn* and *bennen* between 1664-1669 (P1) and 1775-1780 (P2). The distribution of *zijn* and *bennen* had not shifted to a difference considered to be statistically significant: $\chi^2(1)= 1.87$, $p = 0.17$, though the individual probability trends suggest that *zijn* selection increased from time P1 (0.86 probability) to P2 (0.91 probability). Figure 3 summarizes the results.

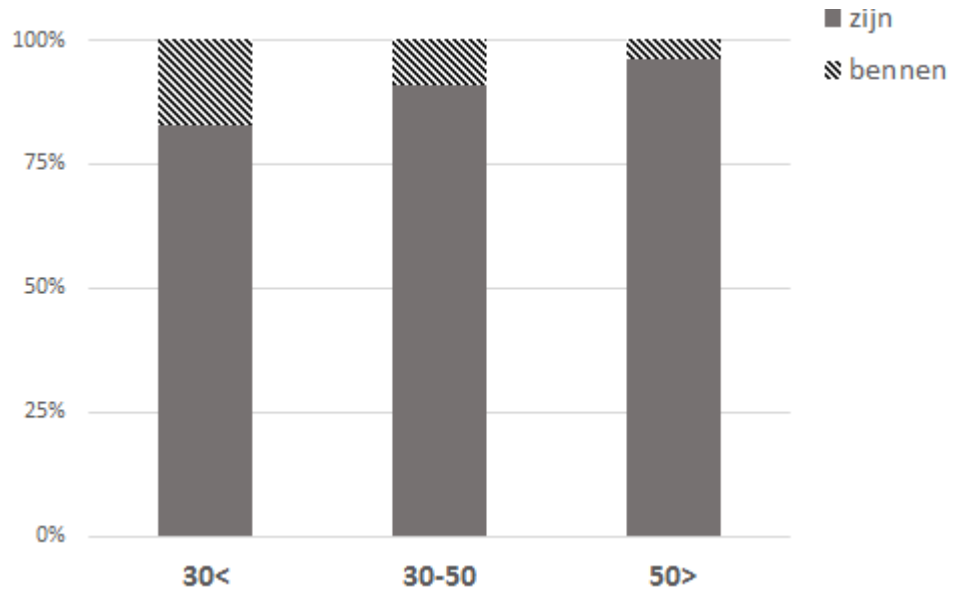

Figure 3. bennen *vs.* zijn *selection across age*

5. Discussion

The variant *zijn*, making up the majority of all frequency counts, had already been firmly established in Early Modern Dutch; the difference in distribution between *bennen* and *zijn* selection in the 17th and 18th centuries is not significant. However, the selection of *zijn* across subject agreement did indicate a significant pattern. As the results in figure 1 show, *zijn* occurred statistically significantly more with the third person plural than it did with the first person plural (there was insufficient data for the second person plural to make conclusions). The third person plural, which has a wider usage scope in discourse (cf. Kuryłowicz 1947; Aalberse and Don 2009: 330, 2011: 346), became unvaried sooner, as opposed to the first person plural, which was more resistant to the loss of variability. Based on these findings, I postulate that the uniformity of the third person plural ultimately spread on to the first and second person plural. However, this must be verified with data from later periods, as the results of the diachronic comparison between the 17th and 18th centuries were inconclusive.

As for social patterning, both social stratification and age indicate a significant patterning. In terms of class, as has been shown in figure 2, the higher one's social status, the higher the likelihood of selecting *zijn*. This suggests that speakers were aware of the variation, and that the *zijn* alternate was presumably linked to prestige. Example (3)

provides an instance of a speaker who corrects herself, further suggesting that speakers were aware of the variation:

(3) *wij* ***bennen*** *wij* ***bennen*** *wij* ***sijn*** *in* *de* *handt* *des* *heeren*
we are we are we are in the hand the-GEN Lord
'We are in the Lord 's hands'
(Magrietje Robbers, November 1664: *Brieven als Buit* corpus)

These facts suggest that the obsolescence of *bennen* was a change from above; as speakers opted for the more prestigious variation *zijn*, *bennen* fell in disuse.

There may also have been a literacy effect involved. Official texts such as the *Statenbijbel* (1637 translation of the Bible in Dutch) utilized the *zijn* alternate for plural person forms. The sentence in (4) provides an illustration.

(4) *Wie* *is* *mijne* *moeder/* *ende* *wie* ***zijn*** *mijne* *broeders*?
Who is my mother and who are my brothers
'Who is my mother and who are my brothers?'
(Matthew 12:48, in the *Statenbijbel* 1637 translation[2])

It is possible that the higher classes were literate at a greater degree than the lower classes, thereby having more exposure to the form *zijn*. A more in-depth study of literacy in the Early Modern Dutch era is needed, if we want to determine the role of *zijn* in official texts had on its usage in spoken language.

As for the age factor, parallel to class, the older the speaker is, the more likely it is for the speaker to use the alternate *zijn*. Literature on age is still scarce in the sub-field of historical sociolinguistics. Blas Arroyo's (2016: 22) study on modal periphrasis, suggests for Early Modern Spanish that youths utilized the innovative variant (a periphrastic modal form with prepositional *de*) much more than adults did (who did not combine the modal with the preposition). Unlike Blas Arroyo's findings for Early Modern Spanish, the trend described here, does not necessarily illustrate the effect of adopting innovative forms over conservative forms, but rather the impact of literacy and standardization over vernacularity. Younger speakers utilize the vernacular at a greater rate than older speakers, who have mostly adopted the prestigious form *zijn*. What this may suggest is that younger speakers did not experience the same social pressures as older speakers, who might have been more conscious of which form to select. It might also have been possible that *bennen* was used to mark in-group identity amongst younger speakers, though this must be researched in more detail.

6. Conclusion and implications

This investigation offered a quantitative sociolinguistic analysis of the *benne* vs. *sijn* plural alternation in Early Modern Dutch. Prior investigations on developments in the ZIJN paradigm had focused primarily on a comparison between Middle Dutch and (Standard) Modern Dutch forms; the literature is mostly silent on ZIJN developments taking place in

[2] *Statenbijbel* (1637), http://www.bijbelsdigitaal.nl/view/, last accessed April 2017.

Early Modern Dutch. Moreover, the historical colloquial variant *benne* has few mentions in the literature; an in-depth discussion in regards to its alternation with *zijn* had not been attempted previously. I attempted to address this gap in the literature by providing a corpus analysis of the *bennen* and *zijn* forms in the 17th century and late 18th centuries using the *Brieven als Buit* corpus. My investigation provided an analysis of ZIJN plural developments both in terms of the linguistic and social factors. The consideration of social factors was especially relevant for the phenomenon discussed here. Class patterning was found to be salient: speakers with a higher social status were the most likely to select the form *zijn*. The fact that *zijn* form was adopted significantly more by the higher classes provides reason to believe that the form *zijn* was linked to prestige (viz. a change from above), suggesting that *bennen* fell in disuse as speakers opted for the prestigious alternative, *zijn*. I further argued that a literacy effect may have influenced the obsolescence of *bennen*, as the *zijn* form is used exclusively in official texts. If it is true that the lower classes had less exposure to literature, it may have influenced the selection of *bennen* and *zijn*. This could be confirmed with future research. Lastly, parallel to the findings for the class index, it was shown that older speakers were more likely to use *zijn* than younger speakers. This pattern may suggest that younger did not experience the same social pressures to conform to the prestigious speech as did older speakers. So far, the investigation here analyzed linguistic and social factors separately, and were not considered in the larger context, that is, the emergence of a standardized language and increased literacy. Future research directions may involve analyzing the interaction between social and linguistic factors and this historical context more deeply.

References

Aalberse, Suzanne, and Jan Don. 2007. Impoverishments in Dutch Dialects. *Online Proceedings of the Mediterranean Morphology Meetings*. Vol. 6.

Aalberse, Suzanne, and Jan Don. 2009. Syncretism in Dutch Dialects. *Morphology* 19: 3-14.

Aalberse, Suzanne, and Jan Don. 2011. Person and Number Syncretisms in Dutch. *Morphology* 21: 327-350.

Blas Arroyo, José Luis. 2016. The rise and fall of a change from below in Early Modern Spanish: the periphrasis *deber de* + infinitive in texts of linguistic immediacy. *Journal of Historical Linguistics* 6(1): 1-31.

Brieven als Buit/*Letters as Loot* corpus. http://brievenalsbuit.inl.nl Leiden University. Compiled by Marijke van der Wal (Programme leader), Gijsbert Rutten, Judith Nobels and Tanja Simons, with the assistance of volunteers of the Leiden-based Wikiscripta Neerlandica transcription project, and lemmatised, tagged and provided with search facilities by the Institute for Dutch Lexicology (INL). 2nd release 2015. Last accessed 22/09/2018.

Chambers, Jack. 2002. Patterns of variation including change. In Jack Chambers, Peter Trudgill, and Natalie Schilling-Estes (Eds.) the Handbook of Language Variation and Change. Oxford: Blackwell. 349-372.

D'Arcy, Alexandra, and Sali Tagliamonte. 2015. Not always variable: Probing the vernacular grammar. *Language Variation and Change* 27, 255-285

Donaldson, Bruce. 1981. *Dutch Reference Grammar*. The Hague: Martinus Nijhoff.

Donaldson, Bruce. 1983. *Dutch: a Linguistic History of Holland and Belgium*. Leiden: Martinus Nijhoff.

Dossena, Marina. 2007. "As This Leaves Me at Present" —Formulaic Usage, Politeness, and Social Proximity in Nineteenth-century Scottish Emigrants' Letters. In: Stephan Elspaß, Nils Langer, and Joachim Scharloth (eds.), *Germanic Language Histories 'From Below' (1700-2000)*. Berlin and New York: De Gruyter. 13-29.

Elspaß, Stephan. 2007. A twofold view 'from below'. New perspectives on language histories and historical grammar. In: Stephan Elspaß, Nils Langer, and Joachim Scharloth (eds.), *Germanic Language Histories 'From Below' (1700-2000)*. Berlin: De Gruyter. 3-9.

Elspaß, Stephan. 2012. The use of private letters and diaries in sociolinguistic investigation. In: Juan Manuel Hernández-Campoy, and Juan Camilo Conde-Silvestre (Eds.), *the Handbook of Historical Sociolinguistics*. Wiley-Blackwell. 156-169.

Frijhof, Willem, and Marijke Spies. 1999. 1650. *Bevochten eendracht*. The Hague: Sdu.

Hock, Hans. 1986. Principles of Historical Linguistics. Berlin: Mouton de Gruyter.

Kielkiewicz-Janowiak, Agnieszka. 2012. Class, Age, and Gender-based Patterns. In: Juan Manuel Hernández-Campoy, and Juan Camilo Conde-Silvestre (eds.), *The Handbook of Historical Sociolinguistics*. Wiley-Blackwell. 307-331.

Koch, Peter, and Wulf Oesterreicher. 1994. Schriftlichkeit und Sprache. Hartmut Günther and Orto Ludwig (Eds.), *Writing and Its Use. An Interdisciplinary Handbook of International Research*. Vol 1. Berlin: Mouton de Gruyter. 587-604.

Koelmans, Leendert. 1978. *Inleiding tot het lezen van zeventiende-eeuws nederlands*. Utrecht: Bohn, Scheltema and Holkema.

Kuryłowicz, Jerzy. 1947. La nature des procès dits analogiques. *Acta Linguistica* 5: 17-34.

Labov, William. 1972. The Social Motivation of a Sound Change (chapter 1). In William Labov (ed.) *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press. 1-41.

Labov, William. 1990. The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2(2):205–254.

Rutten, Gijsbert, and Marijke van der Wal. 2014. *Letters as Loot: a Sociolinguistic Approach to Seventeenth- and Eighteenth-century Dutch*. Amsterdam: John Benjamins.

Statenbijbel, 1637. (1637 Bible translation). http://www.bijbelsdigitaal.nl/view/. Last accessed April 2017.

Toorn van den, Maarten, Wilhelmus Pijnenburg, Jan Leuvensteijn, Joop van der Horst. (1997) *Geschiedenis van de nederlandse taal*. Amsterdam: Amsterdam University Press.

Weinreich, Uriel, William Labov, and Marvin Herzog. 1968. Empirical Foundations for a Theory of Language Change. In: W.P Lehmann and Yakov Malkiel (eds.), *Directions for Historical Linguistics: A Symposium*. Austin: Texas University Press. 97-188.

Zeger, Scott, and Kung-Yee Liang. 1986. Longitudinal Data Analysis for Discrete and Continuous Outcomes. *Biometrics* 42: 121-130.