

From *zijt* to *bent* ‘(you) are’ in Early Modern Dutch: a view “from below” approach

Christina I. Terpstra
University of Calgary

Abstract

Dutch exhibits suppletion in the verb ZIJN ‘be’ with b-roots and s-roots merged into a single paradigm (Donaldson, 1983, p. 182). By the end of the Middle Dutch period (late 16th century), second person singular s-root form zijt and b-root form bent were in competition. Donaldson (1983, p. 182) suggests that the change was motivated by a parallel with the first person singular (present) form ben. However, this account lacks an empirical basis. With the use of the Brieven als Buit (‘Letters as Loot’) corpus, my investigation attempts to address this gap by adopting a “history from below” approach (Elspeß, 2007, p. 5; 2012, p. 160), offering a quantitative analysis on linguistic as well as social factors conditioning this change. The results of the analysis suggest that the shift from zijt to bent occurred at a faster rate in the active, rather than the passive voice, the indicative, rather than the imperative mood, and in predicate rather than auxiliary verb function; these domains are considered more basic with a wider usage (Kuryłowicz, 1947). However, this language change cannot be fully accounted for by a strictly language-internal framework; women selected the bent alternate at a higher rate than men, reinforcing that both social and linguistic factors must be considered to provide a comprehensive account of language change.

Key words: Early Modern Dutch, historical sociolinguistics, morphological change, view “from below”

1 Introduction

At the end of the Middle Dutch era (circa late 16th century), the verb ZIJN ‘be’ had two competing forms for the second person singular, namely conservative *zijt* and innovative *bent*. From the Early Modern Dutch period (circa 17th to 19th century) onwards, the innovative form *bent* came to win out and *zijt* fell into disuse for most varieties of Netherlandic Dutch. In this investigation, I analyze how and why the shift took place, adopting a view “from below” approach, involving texts representative of colloquial language used by men and women of all socio-economic ranks. This approach examines linguistic as well as extra-linguistic factors potentially playing a role in the morphological change of the verb ZIJN. This analysis sampled data from the *Brieven als Buit* (“Letters as Loot”) corpus¹, a private letter corpus representing a diverse authorship in terms of gender and socio-economic class.

The remaining sections are organized as follows: section 2 provides an overview of the literature covering shifts in the second person singular in the ZIJN paradigm. I discuss how prior literature fails to properly account for the *zijt/bent* alternation and elaborate on the motivations behind analyzing the data through the view “from below” approach. Section 3 lays out the methodology. The results and discussion of the analysis are provided in section 4 and 5 respectively. Section 6 offers concluding remarks and suggestions for future research.

2 Second person singular shifts in the ZIJN paradigm

The literature available on the diachronic developments of the ZIJN paradigm are scant. From Old Low Franconian (Old Dutch) to Middle Dutch, the ZIJN paradigm is relatively stable; particularly, no attestable lasting change is observed for the second person singular, which takes the form of *bis(t)* from roughly the 10th- to the 16th century (van den Toorn et al., 1997, p. 41, 58, 118). By the end of the Middle Dutch era, shifts took place in the second person singular. According to Donaldson (1983, p. 182), when the second person singular pronoun *du* ‘you’ disappeared, so did the word-form *bist*. In its place came the pronoun *jij* and the word-form *bent*. Aalberse and Don (2011) state that there is an intermediate stage between Donaldson’s (1983, p. 182) description of the *bist* to *bent* transition. When the form *bist* fell out of use, it left a gap which was temporarily filled by the second person plural form *zijt* (Aalberse & Don, 2011, p. 347-348). A pilot survey of my investigation confirmed the presence of this intermediate stage in the corpus, that is, the *zijt* and *bent* alternation. Instances of the form *bist* were not observed.

As to an explanation for why these shifts take place, Donaldson claims that the *bist* > (*zijt*) > *bent* shift is governed by Kuryłowicz’s (1947) principle of (analogical) leveling, according to which morphologically, syntactically, or semantically related elements become more similar to each other (Hock, 1986, p. 167) –in this context, the

¹ My investigation relied primarily on the following resources: *De Geïntegreerde Taal-bank*, an Old and Middle Dutch dictionary, Kruyskamp’s (1961) edition of the *Van Dale Groot woordenboek den nederlandse taal*, a dictionary on contemporary Dutch, and Verdam’s (1911) *Middel nederlandse hand woordenboek*, a Middle Dutch dictionary, and finally, the *Letters as Loot/Brieven als Buit* corpus (details regarding the corpus follow in section 3).

second person singular *bent* becoming analogous with the first person singular *ben* (Donaldson, 1983, p. 182). The paradigm is constrained by the one-form-one-meaning principle, whereby if a particular morphophonemic alternation does not signal an important meaning contrast, it may be eliminated (Hock, 1986, p. 169). The direction of the leveled paradigm is governed by “basicness”. “Basic” forms constitute 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, p. 183), e.g. the third person is considered to be more basic than the first or second person, the indicative mood to be more basic than the subjunctive or imperative mood, and the active voice to be more basic than the passive voice (Hock, 1986, p. 217-220). However, Donaldson’s (1983) explanations are speculations at best as no empirical evidence is offered for these claims.

Analyzing language change through a strictly system-internal approach as offered by Kuryłowicz (1947) has been criticized. Weinreich et al. (1968, p. 177) argue that by abstracting language from its community matrix, one risks omitting socially conditioned variation from an explanation of linguistic change. Incorporating sociological factors in linguistic analyses has shown to be a highly valuable explanatory tool. For example, Blas Arroyo (2014, p. 22-23) observes that the structure *deber de* + infinitive as opposed to *deber* + infinitive ('have to, must/should') was adopted and spread more by youths than by adults in the 16th century. Incorporating a social lens to diachronic analyses, however, presents a number of challenges; historical sociolinguistic studies must reconstruct the society in addition to the language it embeds. Exclusive literacy and access to written communication by the predominantly upper class, the more careful preservation of these texts, and bias against the use of vernacular texts for historical linguistic analyses has led to the imbalanced representation of the historical registers spoken (Elspaß, 2007, p. 3; 2012, p. 159-161).

The language history “from below” aims to counterweigh the overrepresentation of upper-class registers found in previous research (Elspaß, 2007, p. 5; 2012, p. 160). This approach examines linguistic patterns produced by the marginally-represented majority population which includes farmers, artisans, soldiers, and housemaids (Elspaß, 2007, p. 5). In order to reconstruct an image of the language of the masses, this approach must utilize texts beyond published works, e.g. “ego-documents” such as private letters and diaries (see Elspaß, 2007, p. 5). A view “from below” also constitutes language change occurring below the level of awareness (the types of variation which are usually not linked to prestige) (Labov, 1994, p. 78). Labov states that such a change may be introduced by any social class, but no instances initiated by the higher strata have been reported (1994, p. 78). In conclusion, a view from below would require analyses of texts written by the lower classes for such changes to be brought to light.

In conclusion, because the form *bist* virtually disappears by the end of the Middle Dutch period (Donaldson, 1983, p. 182), and Modern Dutch primarily uses *bent*, the Early Modern Dutch period logically seems to be the locus of the change, and thus serves as the focus-period of my investigation. Moreover, since no prior investigation has provided an empirically-sound analysis of the *zijt* > *bent* shift incorporating potential internal and external factors, providing such an empirically-supported analysis of this shift is the ultimate aim here.

3 Methodology

In order to analyze the *zijt* > *bent* shift, the *Brieven als Buit* ('Letters as Loot') corpus was utilized. This is a corpus of approximately one thousand 17th-to-19th century private letters sent by Dutch men and women from a wide range of ages, geographical and socio-economic backgrounds. This diverse corpus was chosen as it most closely captures the Early Modern Dutch vernacular compared to other types of text (see Koch & Oesterreicher, 1985, 1994, on text type and degree of vernacular representation). The dependent variable is binary; no alternate forms in addition to *zijt* and *bent* were identified in the corpus. Data was collected by searching the corpus for instances (tokens) of the alternating word-forms along with their patterning linguistic and extra-linguistic information. Section 3.1 and 3.2 respectively detail the linguistic and extra-linguistic factors part of the analysis.

3.1 Linguistic variables

Tense/aspect: tokens were recorded along with information about the TAM (Tense/Aspect/Mood) values of the constructions in which they were situated. In terms of morphological tense, Dutch has a simple present and two past tenses (the future is expressed analytically with auxiliaries); the (preterite) simple past tense is made up of word-forms from the *wes-* sub-paradigm of ZIJN (Donaldson 1981, p. 139; 1983, p. 182). Furthermore, Dutch also has an analytic past perfect tense, which takes the structure of ZIJN + *ge-V_{INF}* (Donaldson 1981, p. 141-145). Dutch also has analytic imperfect aspect (Donaldson 1981, p. 165-166) constructions, the most common being ZIJN + *aan_{PREP}* + *het_{DET}* + *V_{INF}* structure (Ebert, 2000; Behrens et al., 2013). In order to perform a distribution comparison on tense and aspect usage with the word-form alternates, the tense and aspect constructions were collapsed into a single variable; the tokens were categorized as either [+past, -present] for past perfect constructions, [-past, +present] for simple present constructions, or [-past, -present] for imperfect aspect constructions (simple past tense was not recorded since the simple past tense does not utilize the verb ZIJN).

Mood: Early Modern Dutch has an indicative, imperative, and traces of a subjunctive mood; the subjunctive mood has gradually become obsolete and only persists in archaic expressions, e.g. *God zij dank* 'God be thanked' (Donaldson 1983, p. 182). In Modern Dutch, the imperative mood form is homophonous with the root *wes-*, although Donaldson mentions that *zijt* was used for the imperative in Early Modern Dutch (1983, p. 182). Tokens were recorded for Tense, Aspect, and Mood for the entire construction, whether or not the information was expressed primarily by the ZIJN verb or by another (main) verb.

Voice: Dutch morpho-syntax has the capability to express passive alongside active meaning through the use of ZIJN as an auxiliary in periphrastic constructions (Donaldson 1981, p. 161-165). I incorporated voice as a variable since it may have an influence on the distribution of conservative and innovative word-forms.

Verb function: In addition to tense, mood, and voice, I also incorporated verb function as a variable. Like English *be*, ZIJN can function as an auxiliary, namely in periphrastic constructions (i.e. past perfect tense and progressive aspect constructions), or it can be used as a predicate verb modifying adjectival, adverbial, or prepositional phrases (e.g. *x is beautiful, late, broken, etc.*, and *x is here, inside, at sea etc.*).

3.2 Extra-linguistic variables

Time: the dates of the letters are based on the dates provided by the letters' authors. Letters which contained the relevant *zijt* and *bent* tokens are divided into two time periods corresponding to the 17th and 18th centuries to measure the rate of change. Period 1 (P1) constitutes letters with *zijt* and *bent* tokens written between 1661 and 1673, and Period 2 (P2) constitutes letters with *zijt* and *bent* tokens written between 1779 and 1783. Both timespans contain an exhaustive collection of tokens available in the corpus. The gap period was decided based on the fact that no tokens were found between these two timeframes.

Gender: I added gender as an external factor as numerous socio-linguistic studies (Rydén & Brorström, 1987: p. 200–207; Labov, 1990; Kytö, 1993; Nevalainen & Raumolin-Brunberg, 2003; Blaxter, 2015) have observed a gendered pattern in linguistic variation. Authors' genders are based on inference of their names (Rutten & van der Wal, 2014, p. 9).

Class: my study follows the socio-economic class system proposed by the developers of the corpus (see Rutten & van der Wal, 2014, p. 9-10). The developers use a four-way class system including the low class (blue-collar workers e.g. sailors, servants, or soldiers), low-middle class (petty bourgeoisie, e.g. small-shop keepers, small-scale craftsmen, or minor officials), high-middle class (large-store keepers, or non-commissioned officers), and a high class (bourgeoisie e.g. wealthy merchants, ship owners, academics, or commissioned officers). The extremely poor and the extremely rich are not represented in the corpus as the former did not have access to written communication, while the latter is omitted because of over-representation in external corpora (Rutten & van der Wal, 2014, p. 10). Female authors were assigned a class label according to their husband's or father's class division.

While scribal intervention², (that is, hired writers for speakers with limited literacy as opposed to self-authorship), could potentially influence representability, my preliminary statistical analysis of authorship effect was shown to be inconclusive (see Rutten & van der Wal, 2014, p. 14 on scribal intervention).

3.3 Data collection and quantitative analysis

In the token collection process, spelling variation is a cause of caution, as variation might express phonological/ morphological distinctions. For these particular forms, spelling variation primarily concerned variation in how the vowel, e.g. *bent*: <*bent*> vs. <*bint*>, and *zijt* <*zijt*> vs. <*zeijt*>, as well as the sibilant in the onset of *zijt* e.g., <*seijt*> vs. <*zeijt*> were represented. These contrasts are purely orthographical and had no influence on the distribution of the dependent variables across linguistic factors. Lastly, token doublets occurring in conjunctive phrases i.e. *zijt* [*x*] and *zijt* [*y*] were only counted as one.

The Generalized Estimated Equation (GEE) (Zeger & Liang, 1986; Liang & Zeger, 1986) test was selected as the appropriate procedure to quantitatively measure the significance of shifts and differences in the distribution patterns of (non-)linguistic

² Dutch literacy in the 17th to 18th century lies on a continuum. Part of the population could neither read nor write, others could only read. Scribes were hired in order to communicate with loved ones overseas (Rutten & van der Wal, 2014, p. 14).

information with the *zijt/bent* alternation, because it is able to handle data in a categorical value format (discrete data in the form of token counts, see Johnson, 2008, p. 159), and can compare samples of uneven sizes, as well as cope with missing data, e.g. unknown social variables (Liang & Zeger, 1986, p. 13). The test was performed using SPSS software.

4 Results

A total of 168 speakers were sampled from the *Brieven als Buit* corpus. The smallest number of tokens provided by a single speaker was 1 token, whereas the largest number of tokens provided by a single speaker was 16 tokens. From these speakers, a total of 318 tokens were collected: 257 tokens (80.8%) of the form *zijt*, and 61 (19.2%) of the form *bent*.

A distribution comparison between the two time periods, 1661-1673 and 1779-1783, was performed. The distribution of *zijt* and *bent* had shifted to a difference large enough to be considered statistically significant: $\chi^2(1) = 4.70$, $p = 0.03$. In Period 1 (P1), the likelihood for speakers to select the innovative form *bent* is 0.16, whereas a century later, in Period 2 (P2), the likelihood double to 0.32. The results are illustrated in Figure 1.

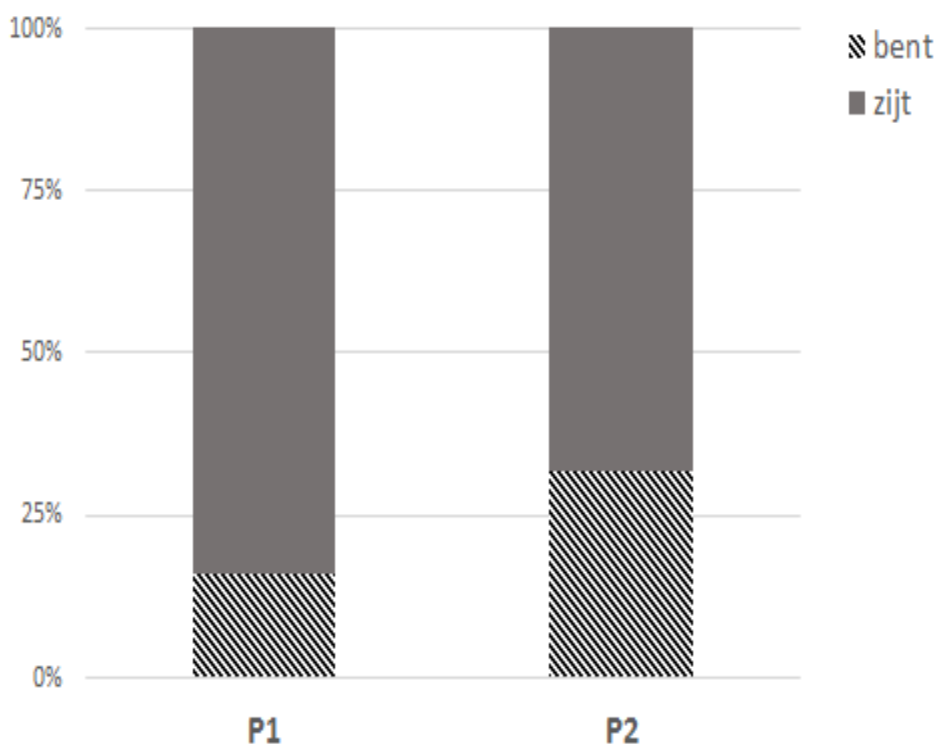


Figure 1: *zijt* vs. *bent* selection across time³

The distribution of *bent* and *zijt* across Tense and Aspect was recorded; three levels were initially considered: the simple present tense, the past perfect tense, and the imperfect aspect. Only two of the three levels provided further grounds for testing; no instances were

³ For each bar graph in this chapter, probability out of 1.0 of selecting the conservative form versus the innovative form was converted into percentages

found of *zijt* or *bent* used in an imperfect aspect context. It is likely that this is due to the fact that occurrences of imperfect constructions are rare in general in the corpus. The present tense and the past perfect tense had enough instantiations of *zijt* and *bent* to be considered for further analyses. There were a total of 278 tokens identified with a present tense context: 232 of the form *zijt* (83.5%), and 46 of the form *bent* (16.5%). As for the past perfect tense, there were a total of 39 tokens identified: there were 24 instances of the form *zijt* (61.5%), and 15 instances of the form *bent* (38.5%). Example (1) provides instances of *zijt* and *bent* in a present tense context, and example (2) provides these in a past perfect tense context.

- 1) a. ghij **zijt** sulcken jongen blom.
 you are such young flower
 (You are such a young flower.)
 (Krijtje Mijnemans, November 1672: Brieven als Buit corpus)⁴
- b. al bent gij uijt min oogen gij en **bent** daerom niet uijt min hart.
 even are you out my eyes you NEG are therefore not out my heart
 (Even if you are out of sight, you are not out of mind.)
 (Magrietje Robbers, November 1664)
- 2) a. gedenk dat gij uijt een eerweerdige stam **zijt** gesproten.
 think that you out a worthy clan are PRF-spring
 (Remember that you come from a noble family.)
 (Hieronymus Sweerts, November 1672: Brieven als Buit corpus)
- b. gij holland geheel vergeete **bent**.
 You Holland completely forget.PRF are
 (You have completely forgotten about Holland.)
 (Gerharda Catharina Wirth-Sluijsken, December 1780)

The distribution difference between the present and past perfect tense was found to be statistically significant: $\chi^2(1) = 5.83$, $p = 0.02$. There is a 0.17 probability for speakers to select *bent* with a present tense context, while the probability of selecting *bent* in a past perfect tense context is more than double: 0.38. The interaction of *zijt* and *bent* selection for the tenses across time was also analyzed. Distribution shifts from P1 to P2 in terms of *zijt* and *bent* selection for present tense/past perfect constructions is significant: $\chi^2(1) = 4.46$, $p = 0.04$. For present tense constructions, the probability of selecting *bent* doubles from P1 (0.14 probability) to P2 (0.29 probability). As for past perfect constructions, *bent* selection likelihood increases from 0.35 probability in P1 to 0.50 probability in P2. Figure 2 summarizes the results in graphic format.

⁴ Unless otherwise specified, all subsequent sample sentences come from the *Brieven als Buit* corpus

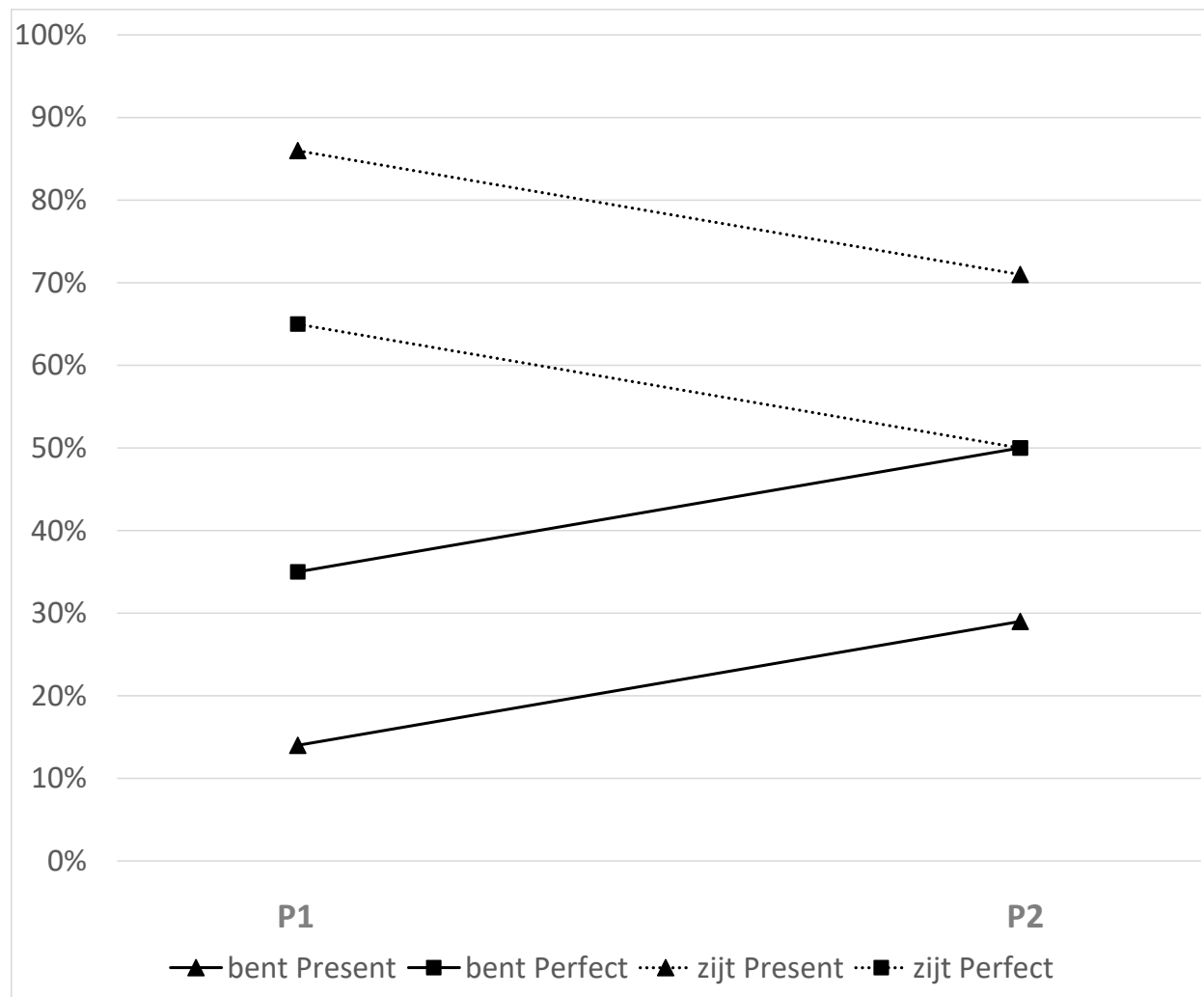


Figure 2: *zigt* vs. *bent* selection in simple present vs. past perfect tense across time

In terms of mood constructions, there are no instances observed in which *bent* was used in contexts other than indicative constructions. As for the *zigt* form, it was found to be selected for indicative as well as imperative constructions. Example (3) illustrates an imperative use of *zigt*.

- 3) **Zigt** versekert.
 be ensured
 (be ensured.)

(Cornelis van Brakel, October 1780)

As mentioned in section 3.1, Modern Standard Dutch uses neither *bent* nor *zigt* to express the imperative, rather, the form *wees* is used (Donaldson, 1983, p. 182). Since *zigt* as an imperative form is eventually overtaken by the contemporary *wees*⁵, a statistical analysis was conducted to determine whether a decrease in *zigt* usage for imperative constructions

⁵ No instances of *wees* were found in the corpus

is observable between the 17th (P1) and 18th century (P2); results indicate, however, that this is not the case: $\chi^2(1) = 0.03$, $p = 0.86$.

In addition to Tense, Aspect, and Mood, Voice was also considered. Tokens were sorted according to whether the construction had active or passive meaning. A total of 162 instances of active voice constructions were identified, of which 109 instances occurred with the *zijt* alternate (67.3%), and 53 instances occurred with the *bent* alternate (32.7%). As for the passive voice, a total of 156 instances were identified, of which 148 constructions occurred with the *zijt* alternate (94.9%), and 8 constructions occurred with the *bent* alternate (5.1%). Example (4a-b) provides instances of *zijt* and *bent* used within active constructions, and example (5a-b) provides instances of *zijt* and *bent* used within passive constructions:

- 4) a. ghy op soo een prykeloose rys **sijt**.
 you on such a dangerous voyage are
 (you are on a dangerous voyage.)
 (Elisabeth Philipse Amelingh, March 1661)
- b. gij mijn kint **bent** in een vreemd lant
 you my child are in a strange land
 (you, my child, are in a foreign country)
 (Adriana van de Sande-Goeree, June 1780)
- 5) a. ghij vande koopman op dort te lossen **sijt** geordeneert
 you from.the salesman on Dordrecht to unload are PRF-order
 (You are ordered by the salesman to unload in Dordrecht)
 (Jan Muessen Ossenweijder, February 1664)
- b. terwijl uwe tog so gesien **bent** bi de Heer governeer
 while you yet well seen are by the Lord Governor
 (Yet, while you are so well received by the Lord Governor)
 (Pieter Kamp, 1780)

The distribution difference between *zijt* and *bent* across active and passive constructions was found to be statistically significant: $\chi^2(1) = 32.43$, $p < 0.01$. While there is a 0.33 probability of *bent* selection in active voice constructions, passive constructions are almost exclusively combined with the *zijt* form (only 0.05 likelihood for *bent* selection). The extent at which *zijt* or *bent* selection differs for passive/active constructions across time is also statistically significant: $\chi^2(1) = 23.57$, $p < 0.01$. In both the active and passive voice constructions, the probability of selecting *bent* increases from Period 1 to Period 2. In active voice constructions, there is a 0.29 probability of *bent* selection in P1, which increases by more than one and a half times (0.48 likelihood) in P2; by this time, there is a nearly equal chance of *bent* or *zijt* selection in active constructions. While *bent* selection is still relatively low by Period 2 in passive voice constructions, the probability of selecting *bent* increases by fivefold from Period 1 (0.03 probability) to period 2 (0.14 probability). Figure 3 summarizes the results in graph format.

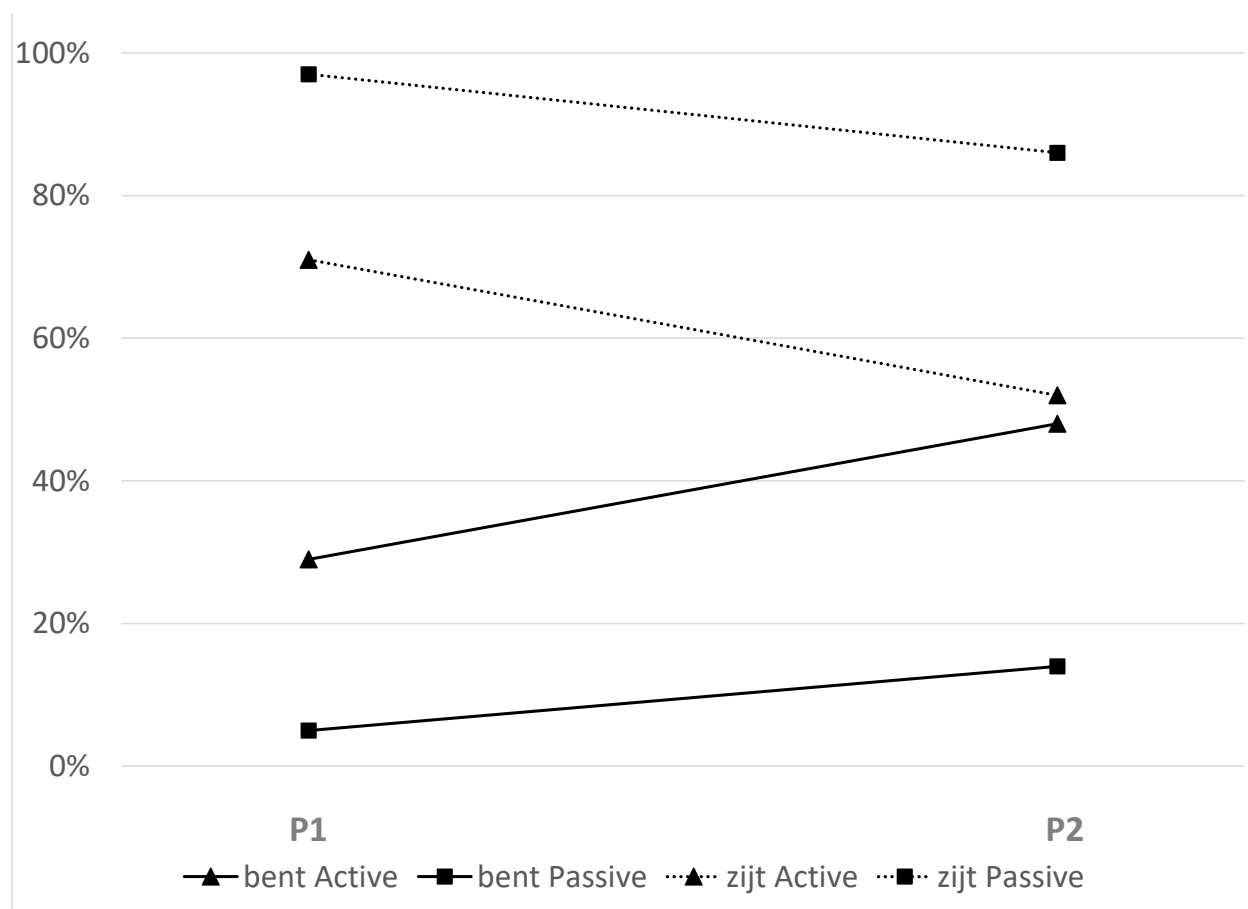


Figure 3: *zijt* vs. *bent* selection in Voice constructions across time

Tokens of *zijt* and *bent* were further sorted according to the function they employed in constructions. As mentioned in section 3, ZIJN can be used either as an auxiliary in periphrastic constructions for the past perfect tense and imperfect aspect, or as a main verb heading a predicate statement (i.e. the head of APs or PPs). An example of each function type is provided in (6):

- 6) a. **sijt** van ons gegroet.
 be from us PTCP-greet
 (Receive greetings from us.)

(Pietertje Fekkes, May 1780)

- b. waer dat gij **zijt**
 where that you are
 (wherever you are)

(Grietje Matijs, September 1664)

A total of 182 tokens were observed with an auxiliary function, 166 of which occurred with the form *zijt* (91.2%), and 16 of which occurred with the form *bent* (8.8%). As for the predicate function, a total of 136 tokens were observed, 91 of which occurred in the form *zijt* (66.9%), and 45 of which occurred in the form *bent* (33.1%). The difference in distribution of *zijt* and *bent* across an auxiliary or predicate function was found to be

statistically significant: $\chi^2(1) = 25.70$, $p < 0.01$; there is a 0.33 probability of *bent* selection for predicate function, while there is only a 0.09 probability of *bent* selection for auxiliary function. Results of the statistical analysis suggest that the difference in distribution of *zijt* and *bent* selection for either an auxiliary usage of predicate usage across time is significant: $\chi^2(1) = 17.18$, $p < 0.01$. In predicate function, the probability of *bent* selection increases by more than one and a half times from Period 1 (0.29 probability) to Period 2 (0.48 probability); there is nearly an equal chance of *zijt* or *bent* selection for predicate constructions in Period 2. While in auxiliary constructions, the probability of *bent* selection is low compared to predicate constructions, the probability of *bent* selection more than doubles between Period 1 (0.07 probability) and Period 2 (0.18 probability). The results of the analyses are summarized in graphic-format in figure 4.

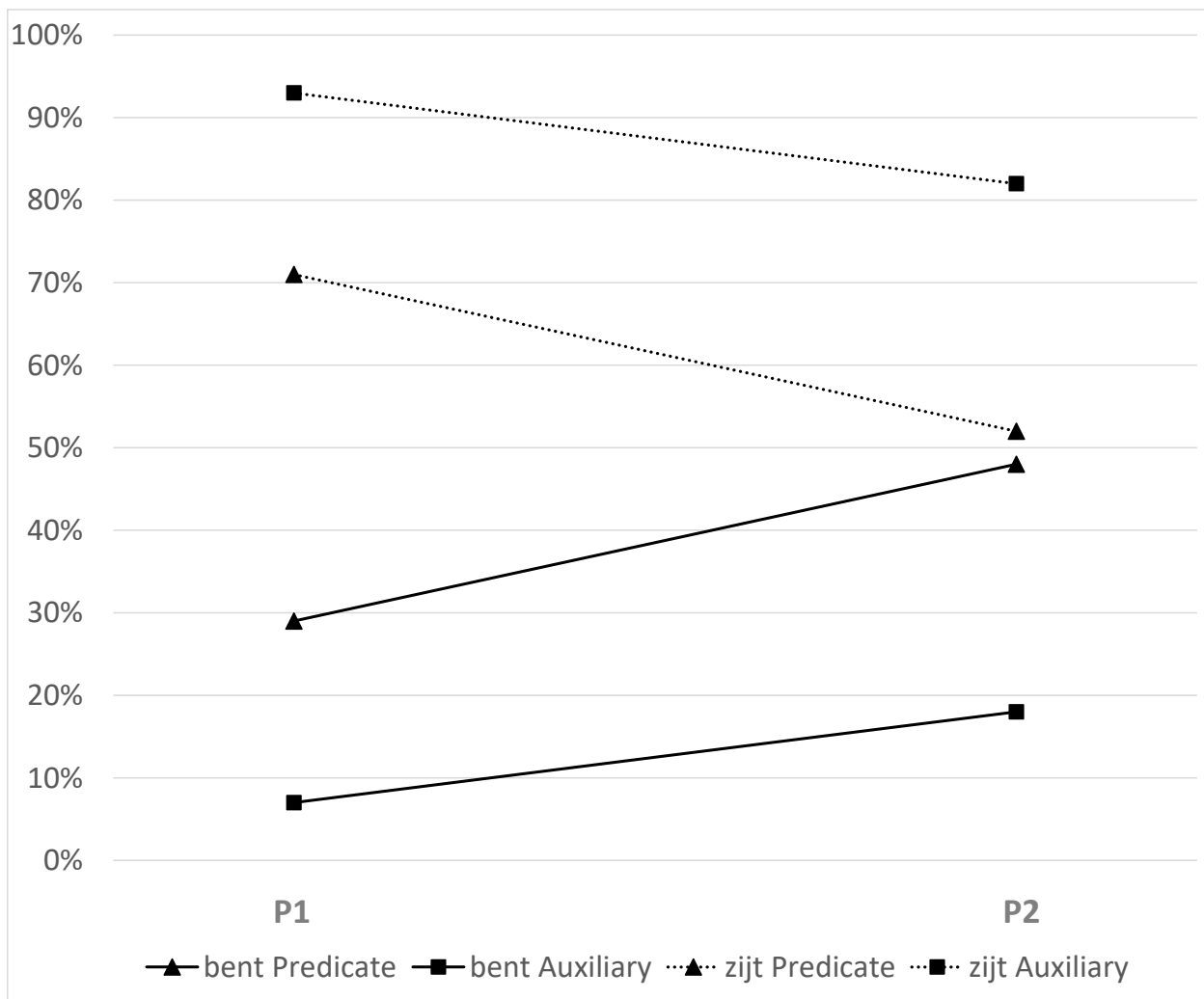


Figure 4: *zijt* vs. *bent* selection in verb function across time

I now turn to the results of the analysis on extra-linguistic factors. To start, a statistical analysis of gender shows that the selection of the innovative or conservative form is gendered: $\chi^2(1) = 5.17$, $p = 0.02$. Women have a 0.25 probability of selecting *bent*, while men only have a probability of 0.13, indicating that women favour the innovative form over the

conservative form to a greater degree than do men. Figure 5 provides a graphic presentation of the inter-gender variation:

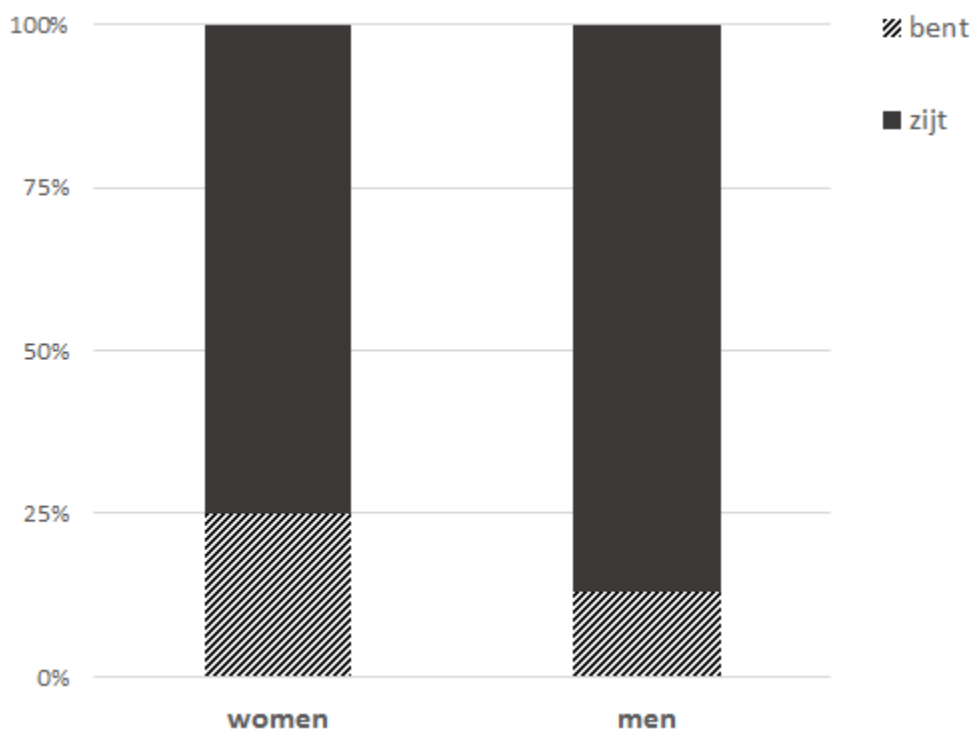


Figure 5: *zijt* vs. *bent* selection across gender

A statistical analysis of class shows that *zijt* or *bent* selection across the various social strata is not significant: $\chi^2(3) = 1.67$, $p = 0.64$. Results from an analysis on *zijt* and *bent* selection for the three age cohorts (<30, 30-50, >50) suggests that the deviations in selection distribution per cohort are not statistically significant either: $\chi^2(2) = 1.53$, $p = 0.47$. In addition to main effects analyses of the social variables, I also performed two-way interaction analyses. The interaction of gender and class does not indicate significant distribution differences in the selection of *zijt* or *bent*: $\chi^2(3) = 1.12$, $p = 0.77$. The interaction of gender and age also does not indicate statistically significant distribution differences in the selection of *zijt* and *bent*: $\chi^2(2) = 4.25$, $p = 0.77$. A class by age interaction analysis could not be performed, due to insufficient data.

5. Discussion

As is apparent from the token frequency, the conservative alternate *zijt* is still well established in the Dutch language between the 17th and 18th century. However, it is evident from the quantitative analyses that I conducted, that the usage of the innovative alternate *bent* had increased: between the 17th and 18th century, the probability of selecting *bent* doubled. Furthermore, the results from my analysis suggest that the shift is conditioned by both linguistic as well as extra-linguistic factors. The gender comparison indicates a statistically significant selection pattern where women lead the trend by

adopting the innovative form at a higher rate than men. This conforms to the trends observed in socio-linguistic literature (e.g. Labov, 2001, p. 321). Since there is no salient socio-economic class patterning, it provides reason to believe that this is change from below; neither form is linked to (non-)prestige, thus it is unlikely that there was awareness on part of the speakers of this shift taking place.

As mentioned in section 2, Donaldson (1983, p. 182) invokes Kuryłowicz's (1947) principle of analogical leveling to account for the *bist* > (*zijt*) > *bent* change. Kuryłowicz argues that alternating forms level to the form given by basic categories, e.g. the third person, active voice, indicative mood, and present tense (Hock 1986, p. 217-220; see also Section 2 of this paper). However, this particular change involving the second person singular does not actually conform to the principle in the most direct way: the *zijt/bent* alternation favours the *b*- root over the *s*- root. The alternation ultimately favours *bent* analogous to the first person form *ben*, not the third person *s*- root form *is*.

While Kuryłowicz's principle of basicness does not hold for person inflection, it does appear that Kuryłowicz sphere of usage principle seems to apply when we consider the selection distribution of the alternates across other domains. For voice constructions, the innovative form *bent* seems to have spread through the active voice domain faster than the passive. This is evident from the fact that the probability of *bent* selection was much higher for active constructions than passive constructions. Diachronically, the findings suggest that after the innovative form became quite well established in active constructions (where it had nearly 50/50 selection probability by the late 18th century), passive constructions eventually followed suit; the probability of speakers using *bent* increased fivefold by the late 18th century. In parallel to the passive voice, the usage of *zijt* for the imperative mood is quite resistant to change as well. Between the 17th and 18th century, there was no statistically significant shift in distribution observed; the loss of *zijt* as an imperative form is likely to be a much more recent development.

These combined findings allude to Kuryłowicz's (1947) principle of basicness; loss of variation happens at a higher/faster degree within the more basic categories. Conversely, less basic categories such as the passive voice and imperative mood are much more resistant to change. Verb function seems to behave in a similar fashion as well. While Kuryłowicz did not conceptualize the basicness of verbal function types, it could be argued that the predicate function is more basic than the auxiliary function, since they have a greater sphere of usage (i.e. occur with non-verb phrases such as adjectival and prepositional phrases). In the 17th century, a predicate usage of *bent* was quite high relative to the auxiliary usage. Then by the late 18th century, the usage of *bent* with auxiliaries caught up with that of the predicate; the probability of *bent* selection for auxiliary usage had nearly tripled.

The only result which seems initially contradictory to the rest of the results are those from the comparison of *zijt* and *bent* usage with the present and past perfect tense. The results indicate that the usage of *bent* with the present tense is relatively low in comparison to the past perfect tense. If predicate constructions (which favour *bent* usage) occurred mostly with the present tense, and auxiliary constructions were mostly recorded for past perfect tense constructions, then why do we observe the opposite trend? In order to proceed with a statistical analysis, the SPSS program requires that all cells of the dataset are filled in. By default, I assigned all instances of imperative and passive constructions with present tense. I suspect that since imperative constructions and most of the passive

constructions were coded with the present tense, this may have counterbalanced the effects of *zijt/bent* selection for the present tense and past perfect tense. Future statistical analyses controlling for these factors may confirm this. Furthermore, it is not clear whether Kuryłowicz's (1947) principle of basicness plays a role here. The present tense is argued to be the pivot for all other tenses (Hock, 1986, p. 217-220). In this case, this seems to be irrelevant since the simple past tense inflections are based on the *wes-* root, which at least as is observed from contemporary usage (Donaldson, 1981, p. 139), does not level to either an *s-* or *b-* root. Likewise, it is unclear to me how basicness intersects with periphrastic tenses.

6 Concluding remarks

In this investigation, I examined language-internal and external factors governing the direction and pace of the shift from *zijt* to *bent* in the Dutch verb *ZIJN*. The phenomenon did not pattern with socio-economic class in a statistically significant way because neither of the forms carried a prestige value. This makes it likely to conclude that this is a change from below speakers' awareness. Without the speakers consciously operating on the alternation, the explanatory burden falls primarily on linguistic factors. This seems to be the case here as changes occurred at a faster pace across the domains of active voice, indicative mood, and predicate function, alluding to the idea proposed by Kuryłowicz (1947) that these categories are more basic.

However, whether Kuryłowicz's (1947) notion of basicness best accounts for the change presented in this analysis is questionable. It is not clear from this principle if the variants ought to be compartmentalized across the various domains monolithically, or if the distribution of this alternation, or any alternations for that matter, transcend category boundaries. The latter would arguably make invoking this principle problematic, yet this is what we find in this analysis –the alternates are found in varying ratios across all of the basic and non-basic domains. To assume the grammar blindly assigns the pivot of leveling according to what form is embedded in a particular basic domain is not empirically supported.

Rather, focus needs to be diverted from the mechanics of the system and brought towards the usage of this system by the speakers. While class did not pattern with the distribution of the alternation, gender did, reinforcing the argument put forth by Weinreich et al. (1968) that linguistic changes cannot solely be accounted for from a system-internal framework. In regards to future research, I propose to examine more closely the interaction between speakers and the system through a Usage-based Grammar framework. The current investigation examined linguistic and social factors as discrete forces and did not account for the interaction between the speakers and the mechanics of the system. One of the main tenets of the Usage-based Grammar approach is that frequency effects influence every domain of grammar (see Barlow & Kemmer, 2000; Bybee & Hopper, 2001). Given the results of this analysis, it seems more likely that it is the interaction of usage on the domains that informs the output of a change, and not the domain itself.

In conclusion, I investigated the *zijt > bent* ('you) are' shift in Early Modern Dutch (17th to 18th century), a linguistic shift for which an empirically-based analysis had not

been previously offered. In addition, the results of my analysis reinforce the importance of considering both linguistic as well as social factors in historical linguistic analyses.

References

- Barlow, M., & Kemmer, S. (Eds.). (2000). *Usage-based models of language*. Stanford: CSLI.
- Behrens, B., Flecken, M., & Carroll, M. (2013). Progressive attraction: On the use and grammaticalization of progressive aspect in Dutch, Norwegian, and German. *Journal of Germanic linguistics*, 25(2), 95-136.
- Blas Arroyo, J. L. (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.
- Blaxter, T. (2015). Gender and language change in Old Norse sentential negatives. *Language Variation and Change* 27, 349-375.
- Bybee, J., & Hopper, P. (Eds.). (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.
- Donaldson, B. (1981). *Dutch Reference Grammar*. The Hague: Martinus Nijhof.
- Donaldson, B. (1983). *Dutch: a Linguistic History of Holland and Belgium*. Leiden: Martinus Nijhoff.
- Ebert, K. (2000). Progressive markers in Germanic languages. In Ö. Dahl (Ed.) *Tense and Aspect in the Languages of Europe*. Berlin: Mouton de Gruyter. 605-653.
- Elspaß, S. (2007). A Twofold View 'From Below'. New Perspectives on Language Histories and Historical Grammar. S. Elspaß, N. Langer, & J. Scharloth (Eds.), *Germanic Language Histories 'From Below' (1700-2000)*. Berlin: De Gruyter. 3-9.
- Elspaß, S. (2012). The Use of Private Letters and Diaries in Sociolinguistic Investigation. J. M. Hernández-Campoy, & J. C. Conde-Silvestre (Eds.), *the Handbook of Historical Sociolinguistics*. Wiley-Blackwell. 156-169.
- Geïntegreerde Taal-bank, de*. (2007-2010). Instituut voor de Nederlandse taal. <http://gtb.inl.nl/?owner=ONW>. Last accessed: April 2017.
- Hock, H. (1986) *Principles of Historical Linguistics*. Berlin: Mouton de Gruyter.
- Johnson, K. (2008) *Quantitative Methods in Linguistics*. Oxford: Blackwell.
- Koch, P., & Oesterreicher, W. (1985) Sprache der Nähe - Sprache der Distanz. Mündlichkeit und Schriftlichkeit und Sprachgeschichte. *Romantisches Jahrbuch* 36, 15-43.
- Koch, P., & Oesterreicher, W. (1994) Schriftlichkeit und Sprache. H. Günther & O. Ludwig (Eds.), *Writing and Its Use. An Interdisciplinary Handbook of International Research*. Vol 1. Berlin: Mouton de Gruyter. 587-604.
- Kruyskamp, C. (1961). *Van Dale Groot woordenboek der nederlandse taal*. The Hague: Martinus Nijhoff.
- Kuryłowicz, J. (1947) La Nature des procès dits analogiques. *Acta Linguistica* 5, 17-34.
- Kytö, M. (1993). Third-person present singular verb inflection in early British and American English. *Language Variation and Change* 5(2), 113-139.
- Labov, W. (1990). The intersection of sex and social class in the course of linguistic change. *Language Variation and Change* 2(2), 205-254.
- Labov, W. (1994) *Principles of Linguistic Change, Vol 1: Internal Factors*. Oxford: Blackwell.
- Labov, W. (2001). *Principles of Linguistic Change, Vol 2: Social Factors*. Oxford: Blackwell.
- The Letters as Loot/ Brieven als Buit corpus*. Leiden University. Compiled by Wal, M., van der, Rutten, G., Nobels, J., & Simons, T., 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. <http://brievenalsbuit.inl.nl> accessed August 2018.
- Liang, K.-Y., & Zeger, S. (1986). Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika* 73, 13-22.
- Nevalainen, T., & Raumolin-Brunberg, H. (2003). *Historical sociolinguistics: Language change in Tudor and Stuart England*. Longman Linguistics Library. Harlow: Pearson Education Limited.
- Rutten, G., & Wal, M., van der. (2014) *Letters as Loot: A Sociolinguistic Approach to Seventeenth- and Eighteenth-century Dutch*. Amsterdam: John Benjamins.

- Rydén, M., & Brorström, S. (1987). The be/have variation with intransitives in English. With special reference to the Late Modern Period. *Acta Universitatis Stockholmiensis: Stockholm Studies in English* 70. Stockholm: Almqvist & Wiksell International
- Toorn, M.C., van den, Pijnenburg, W., Leuvensteijn, J.A., van, & Horst, J.M., van der. (1997) *Geschiedenis van de nederlandse taal*. Amsterdam: Amsterdam University Press.
- Verdam, J. (1911). *Middel nederlandsche hand woordenboek*. The Hague: Martinus Nijhoff.
- Weinreich, U., Labov, W., & Herzog, M. (1968) Empirical Foundations for a Theory of Language Change. W.P. Lehmann, & Y. Malkiel (Eds.), *Directions for Historical Linguistics: A Symposium*. Austin: Texas University Press. 97-188.
- Zeger, S., & Liang, K.-Y. (1986) Longitudinal Data Analysis for Discrete and Continuous Outcomes. *Biometrics* 42, 121-130.

Contact information

Christina I. Terpstra
citerpst <at>ucalgary<dot>ca
School of Languages, Linguistics, Literatures, and Cultures
University of Calgary
2500 University Drive N.W.
Calgary, Alberta, T2N 1N4
Canada