The Loss of Negative Concord: Internal Factors

Amel Kallef

Abstract

This study re-addresses the loss of Negative Concord (NC) and provides evidence for the failure of certain external factors such as prescriptivism to account for the observed decline and ultimate disappearance of NC in Standard English. A detailed study of negation in Late Middle and Early Modern English reveals that the loss of NC was a case of a natural change, i.e. a change triggered by some internal factors rather than the result of some external ones. This change was preceded by a period of variation for which S-curves for all the contexts studied were obtained. A close study of n-words like nothing, nobody, etc. in negative contexts and their ultimate replacement with Negative Polarity Items (NPIs) like anything, anybody, etc. in a number of grammatical environments shows that the decline of NC follows the same pattern across contexts in a form of parallel curvature, which indicates that the loss of NC is a natural change. However, this study reveals that the decline is not constant across time and thus the Constant Rate Hypothesis (Kroch 1989) does not, in that respect, provide an adequate model for this change. Context behaviour suggests an alternative principle of linguistic change, the Context Constancy Principle. A context constancy effect is obtained across all contexts indicating that the loss of NC is triggered by a change in a single underlying parameter setting. Accordingly, a theory-internal explanation is suggested. N-indefinites underwent a lexical reanalysis whereby they acquired a new grammatical feature [+Neg] and were thus reinterpreted as negative quantifiers, rather than NPIs. This lexical reanalysis was triggered by the ambiguous status of n-words between [+Neg] and thus between single and double negative meanings.

*An earlier version of this paper was presented at the Third York-Holland Symposium on the History of English Syntax. I'd like to thank the audience at this conference as well as the editors of this volume for their helpful suggestions and comments. All remaining errors are mine. Correspondence: amellkallef@hotmail.com

1 Introduction

While Modern Standard English (MSE) (1800-1920) is not characterised by the operation of negative concord (NC), in some dialects of English and in certain older forms of the language the operation of negative concord is much stronger. MSE exhibits a virtually uniform [NC] system, while earlier forms of English are characterised by the phenomenon of multiple negation. References to this effect may be found in most general studies of Middle English (ME, 1100-1500), as well as in those of Early Modern English (EMODE, 1500-1800) (e.g. Burnley 1983: 61; Barber 1997: 283). These periods exhibited variable use of [+NC] and [-NC] systems. The following two examples illustrate the [-NC] and [+NC] systems, respectively:

(1) He didn’t hurt anybody.

(2) John didn’t hurt nobody.

It has largely been assumed that the loss of NC was the outcome of prescriptive views on language use, and of taking Latin, a [-NC] language, as a model for the English grammar. Because of these assumptions, the issue of why NC was lost in MSE has not been given enough attention. Therefore, this issue will be re-addressed within the framework of a detailed study of the process of decline of NC in order to find out about the nature of this change which took place in the EMODE period.

2 Negation in Pre-modern English

After the loss of ne, which was the primary negator in Old English and for some of the Middle English period, it became very common to express negation through the use of the secondary negator not together with another negative element such as no, hence the name NC, as the following examples illustrate.

(3) I would not for no good...
   (The Lisle Letters, Vol. V: 305)

(4) I am not able to deserve with no power.
   (The Lisle Letters, Vol. V: 196)

1NC has survived in some English dialects, such as the African-American Vernacular English (Labov 1972: 785), as the following example illustrates:

(6) Ain’t nobody ever thought about pickin’ up nothin’.

2This study excludes the non-standard varieties of English which exhibit an NC system.

3Although Nevalainen (1999) and Mazon (1992) state that the loss of NC was almost completed by the time prescriptive schools existed.

By the Late Middle and Early Modern English periods, which in this study stand for the period running from 1450-1600, speakers had an alternative option: the use of any-words in contexts where n-words were previously used. Competition between these two variants, n-words and any-words, arose. There was a period of variation whereby both grammars, [+NC] and [-NC], coexisted. The notion of coexisting grammars, which implies that the same rate of change in different surface contexts reflects a single underlying parameter change, is supported by evidence in the literature. This is known as Kroch’s (1989, 1994) Constant Rate Effect (CRE). It implies that when one grammar is replaced by another over time, usage frequencies change at the same rate, but not necessarily at the same time. The following section will provide a fuller discussion of the Constant Rate Hypothesis (CRH).

3 The Constant Rate Hypothesis (Kroch 1989, 1994)

Kroch (1989, 1994) argues that, when one grammatical option replaces another with which it is in competition across a set of linguistic contexts, the rate of replacement is the same in all of them, i.e. they do not differ in the rate at which the new form spreads. This is known as the Constant Rate Effect of syntactic change, whereby innovations advance at the same rate across linguistic contexts. According to Kroch (1994), the grammatical analysis that defines the contexts of a change is quite abstract. He suggests that the set of contexts that change together is not defined by the sharing of a certain surface property, but rather by a shared syntactic structure, whose existence can only be the product of an abstract grammatical analysis on the part of the speakers. Thus, the change in use of a particular form in different contexts proceeds at the same rate in all contexts and the loss of one grammatical option should occur at the same rate as the gain in use of its morphosyntactic competitor. The CRH predicts that if the forms are in morphosyntactic competition, the logistic regression lines will be parallel, i.e. they have the same slope - τ - which stands for the rate of change. This contradicts Bailey’s (1973: 79-82) second principle of linguistic change which holds that the new form will be used at different frequencies depending on contexts, i.e., we are more likely to come across the new form in higher frequencies in more favourable contexts than in others.

We now turn to Bailey’s (1973: 77) first principle of linguistic change which holds that linguistic changes follow an S-shaped curve. This principle seems to reflect a characteristic property of changes that have been studied quantitatively. In fact, the idea was also supported in work by Osgood and Sebeok (1954); Weinreich, Labov, and Herzog (1968); Kroch (1989) and Chambers (1992), among many others.
4 Data sources and categorisation

This study makes use of a corpus of Late Middle English and Early Modern English texts created for the purpose of this study. The texts selected range in date from the second half of the fifteenth century until the end of the sixteenth century, i.e. 1450 up to 1599. Only private correspondence texts were used as a source of data. All the texts belonging into the time span between 1450 and 1599 are included and all occurrences of n-words and any-words within these texts are counted. All private letters that belong to our period based on their dates of writing established in the printed edition were included (see Appendix A). Translated letters and letters which have a quasi-public nature, namely those by and to members of the Royal Family, were excluded from our analysis.

We have looked at all negative constructions in the first place and collected only those that make use of two negative elements within a single clause or across clause boundaries, and utterances that make use of negative polarity items preceded by a negative element, either the sentential negator not or a n-item. These cases were categorised in terms of grammatical constructions, namely non-coordinate (examples (5)-(6)) and coordinate (examples (7)-(8)) negative environments.

(5) 'That ye wryt to me no letters.' (The Cely Letters: 10)
(6) 'That I shulde not take any writyng of any inyn.' (Christ Church Letters: 66)
(7) 'I had none, ner he deluyerd me none.' (The Cely Letters: 18)
(8) 'Ne bounds for any trett of your seid.' (Christ Church Letters: 41)

They were also categorised in terms of two grammatical functions, namely Objects and Adjuncts, as they occur within these two grammatical constructions. The examples (9) and (10) illustrate Objects as they occur in both non-coordinate and coordinate contexts. Examples (11)-(12) illustrate Adjuncts as they occur in both these constructions.

(9) 'That ye wryt to me no letters.' (The Cely Letters: 10)
(10) 'Nor send me no star. [sterling] money.' (The Cely Letters: 12)
(11) 'Wol natt dyshease yow off youre houase no longer.' (The Cely Letters: 170)
(12) 'Ne forthere prosede in no seche materie.' (The Paston Letters: 82 & 137)

In other words, all instances of n-items in positions where they c-commanded all instances of n-items and corresponding NPIs occurring in object and adjunct grammatical functions within non-coordinate and coordinate constructions are included. For the purpose of this study, only cases of NPIs in negative clauses were collected; those within conditional and interrogative clauses were excluded, as they are irrelevant to our research. Our negative contexts can then be summarised as follows:

\[
\frac{e^{k+s}}{1 + e^{k+s}}
\]

4 NEG refers to the main sentential negator whether it is not or any of its earlier forms and spellings, such as negit, nouth, net, nate, etc.
5 N-items refer to terms such as nothing, none, never, etc. in their various spellings. There is, however, an issue over whether n-words are indefinite or negative quantifiers (Zammitani 1993).
6 Stage 1 (1450-1474); Stage 2 (1475-1499); Stage 3 (1500-1524); Stage 4 (1525-1549); Stage 5 (1550-1574); Stage 6 (1575-1699). For details of the texts within each period see Appendix B.
7 The goodness of fit can only be approximated within the statistical model.
Amel Kallel

The logit, i.e. the logistic transform, in equation (16) is an equivalent form of equation (15) but produces a straight line, standing for a linear function of time. s is the slope of this line. k is the intercept, and is related to the frequency of the old/new form at some fixed point in time.

\[ \text{Logit} = \ln \frac{p}{1-p} = k + st \]

The change in the frequencies of these two grammatical alternatives is interpreted according to the Constant Rate Hypothesis (CRH) in an attempt to find out about the rate of replacement of one form by another in some grammatical environments. This is crucial as it reveals issues related to the nature of the change: whether the change is the outcome of competition between grammatically incompatible options, and thus whether the observed changes in the surface structures are triggered by a change in a single underlying parameter setting.

6 Data analysis

In this section we analyse the shift in the percentages of use of NC in non-coordinate contexts (Table 1). In stage 1, 1450-1474, we record 105 cases of NC, equal to 83.3% out of a total of 126 data points. This percentage gradually declined in our period until NC virtually disappeared at stage 5 with 31.1%, and went completely out of use at stage 6, the last quarter of the sixteenth century, with only 0.6%. This decline in the use of NC in non-coordinate contexts meant a corresponding gradual increase in the use of negative polarity items in contexts where n-words used to be used. In stage 1, negative polarity items were used only 16.7% of the time; this frequency, however, rose to 96.9% in stage 5 and the use of NPIs in negative contexts was fully established in stage 6 with a percentage of 99.4%.

The plot of the transformed data given in Figure 1 shows the observed probabilities fitted into the logistic regression model. The decline of NC in non-coordinate constructions follows an S-shaped curve. The S-shaped curve, a general characteristic of linguistic changes suggests that the change follows the slow-quick-slow pattern. Table 2 summarises the frequencies of use of NC and NPIs in coordinate contexts throughout the six stages.

<table>
<thead>
<tr>
<th>Stage</th>
<th>NC</th>
<th>NPI</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>stage 1</td>
<td>105</td>
<td>83</td>
<td>21</td>
</tr>
<tr>
<td>stage 2</td>
<td>91</td>
<td>78.4</td>
<td>25</td>
</tr>
<tr>
<td>stage 3</td>
<td>15</td>
<td>48.1</td>
<td>14</td>
</tr>
<tr>
<td>stage 4</td>
<td>81</td>
<td>43.3</td>
<td>105</td>
</tr>
<tr>
<td>stage 5</td>
<td>5</td>
<td>3.1</td>
<td>155</td>
</tr>
<tr>
<td>stage 6</td>
<td>1</td>
<td>0.6</td>
<td>156</td>
</tr>
</tbody>
</table>

Table 1: The frequency of n-words and NPIs in non-coordinate constructions by stage.

Figure 1: The observed data for non-coordinate contexts plotted against the fitted logistic regression.

---

Non-coordinate constructions will be referred to as type 1 in the graphs and tables; coordinate constructions as type 2.

By 'period' we will be referring to the span of 150 years that we studied; stages 1-6 are spans of 25 years each.

For reasons of space, only key graphs will be presented in this paper.
Amel Kallel

The picture of the change is very similar to the one observed in non-coordinate contexts. Again, we notice that overall there is an obvious decline of NC in coordinate contexts. The frequency of use of NC in coordinate contexts dropped from 96.5% in stage 1 to 7.0% in stage 6. The change again reflects the ongoing rise in the frequency of negative polarity items in contexts where n-words were used. The corresponding percentage in NPIs rose from only 3.5% in stage 1 to 93.0% in stage 6. Figure 2 indicates that the S-shaped curve also obtains for coordinate contexts.

<table>
<thead>
<tr>
<th>Type 2: Coordinate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage</td>
</tr>
<tr>
<td>N %</td>
</tr>
<tr>
<td>stage 1</td>
</tr>
<tr>
<td>stage 2</td>
</tr>
<tr>
<td>stage 3</td>
</tr>
<tr>
<td>stage 4</td>
</tr>
<tr>
<td>stage 5</td>
</tr>
<tr>
<td>stage 6</td>
</tr>
</tbody>
</table>

Table 2: The frequency of n-words and NPIs in coordinate constructions by stage.

Figure 2: The observed data for coordinate contexts plotted against the fitted logistic regression.

An analysis of the process of loss of NC in the two grammatical functions, objects and adjectives, as they occur within the two grammatical constructions yields the same results: an S-shaped curve (see Appendix C).\(^1\)

\(^1\)Initially, all grammatical functions, such as subjects, embedded subjects, noun post-modifiers, etc. were considered, but only objects and adjectives were analyzed statistically, as the others were low in absolute frequency.

\(^2\)Objects will be referred to as Function 1 in the graphs and tables; Adjectives as Function 2.

The Loss of Negative Concord: Internal Factors

<table>
<thead>
<tr>
<th>OBJECTS (FUNCTION 1): Non-Coordinate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage</td>
</tr>
<tr>
<td>N %</td>
</tr>
<tr>
<td>stage 1</td>
</tr>
<tr>
<td>stage 2</td>
</tr>
<tr>
<td>stage 3</td>
</tr>
<tr>
<td>stage 4</td>
</tr>
<tr>
<td>stage 5</td>
</tr>
<tr>
<td>stage 6</td>
</tr>
</tbody>
</table>

Table 3: The frequency of n-words and NPIs in Objects in non-coordinate contexts by stage.

<table>
<thead>
<tr>
<th>OBJECTS (FUNCTION 1): Coordinate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage</td>
</tr>
<tr>
<td>N %</td>
</tr>
<tr>
<td>stage 1</td>
</tr>
<tr>
<td>stage 2</td>
</tr>
<tr>
<td>stage 3</td>
</tr>
<tr>
<td>stage 4</td>
</tr>
<tr>
<td>stage 5</td>
</tr>
<tr>
<td>stage 6</td>
</tr>
</tbody>
</table>

Table 4: The frequency of n-words and NPIs in Objects in coordinate contexts by stage.

Tables 3 and 4 record the occurrences and percentages of both NC and NPIs in negative clauses in objects in non-coordinate and coordinate contexts. As in the case of other contexts, there is an overall decline of n-words in coordinate constructions corresponding to a rise in the frequency of polarity items, across time. NC in objects was used with a frequency of 80.0% in non-coordinate contexts and 95.0% in coordinate contexts at stage 1 and dropped to 1.3% and 3.6% at stage 6 in both grammatical constructions, respectively.

The change in NC in Adjuncts (see Tables 5 and 6) follows the same pattern. We notice an ongoing decline in the percentages of use of n-words as opposed to an ongoing rise in the percentages of use of negative polarity items and an S-shaped curve was also obtained for these contexts.

We would now like to test the Constant Rate Hypothesis and find out whether the decline of NC occurred at the same rate in all studied contexts and whether it took place at the same rate across the time line. Accordingly, linear logistic regression models are used to model the data. We begin by studying the Constant Rate Effect in macro contexts, i.e. non-coordinate and coordinate contexts, and then proceed to the study of micro contexts, i.e. object and adjunct grammatical functions (GPs).
ADJUNCTS (FUNCTION 2): Non-Coordinate

<table>
<thead>
<tr>
<th>Stage</th>
<th>NC</th>
<th>%</th>
<th>NPI</th>
<th>%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>stage 1</td>
<td>59</td>
<td>86.8</td>
<td>9</td>
<td>13.2</td>
<td>68</td>
</tr>
<tr>
<td>stage 2</td>
<td>40</td>
<td>77.0</td>
<td>12</td>
<td>23</td>
<td>52</td>
</tr>
<tr>
<td>stage 3</td>
<td>7</td>
<td>50.0</td>
<td>7</td>
<td>50</td>
<td>14</td>
</tr>
<tr>
<td>stage 4</td>
<td>54</td>
<td>49.5</td>
<td>55</td>
<td>50.5</td>
<td>109</td>
</tr>
<tr>
<td>stage 5</td>
<td>5</td>
<td>1.5</td>
<td>76</td>
<td>98.7</td>
<td>77</td>
</tr>
<tr>
<td>stage 6</td>
<td>0</td>
<td>0</td>
<td>82</td>
<td>100</td>
<td>82</td>
</tr>
</tbody>
</table>

Table 5: The frequency of n-words and NPIs in Adjuncts in non-coordinate contexts by stage.

ADJUNCTS (FUNCTION 2): Coordinate

<table>
<thead>
<tr>
<th>Stage</th>
<th>NC</th>
<th>%</th>
<th>NPI</th>
<th>%</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>stage 1</td>
<td>30</td>
<td>96.8</td>
<td>1</td>
<td>3.2</td>
<td>31</td>
</tr>
<tr>
<td>stage 2</td>
<td>17</td>
<td>89.5</td>
<td>2</td>
<td>10.5</td>
<td>19</td>
</tr>
<tr>
<td>stage 3</td>
<td>8</td>
<td>87.5</td>
<td>4</td>
<td>12.5</td>
<td>12</td>
</tr>
<tr>
<td>stage 4</td>
<td>28</td>
<td>54.4</td>
<td>21</td>
<td>45.6</td>
<td>46</td>
</tr>
<tr>
<td>stage 5</td>
<td>2</td>
<td>12.5</td>
<td>14</td>
<td>67.5</td>
<td>16</td>
</tr>
<tr>
<td>stage 6</td>
<td>1</td>
<td>7.2</td>
<td>13</td>
<td>92.6</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 6: The frequency of n-words and NPIs in Adjuncts in coordinate contexts by stage.

6.1 Macro contexts: grammatical constructions

Without a mathematical model of the $S$-curves, it would be difficult to evaluate certain predictions, as it would not be clear how to measure their rates of change. Visual inspection of the figures given above might suggest that the rates of change for each context are different. However, on examining the curves describing the replacement of one form by another, we notice that this is a wrong interpretation. The logistic function is used to estimate the rates of decline of NC in different contexts in a more accurate way. What the Constant Rate Hypothesis tells us is that if we transform the percentage data and obtain an estimate of the parameters of each curve by fitting the transformed data to a regression line, the slopes of all the lines should be equal. Such a relationship corresponds to the hypothesis that the processing effects on the frequency of the new form in different environments are constant across time. The results of this statistical model are presented below in Table 7.

Figures 3 and 4 show that the lines are parallel and accordingly that the rate of change in both non-coordinate and coordinate contexts is the same. When measured, the slope $s$, which stands for the rate of change, is the same for both constructions and is equal to $-1.126$ units. These graphs indicate that there is a slight overall difference between the two construction types with coordinate constructions (type 2) being higher on the logit scale by an estimated difference of 0.979 units. This difference observed on the logit scale indicates that it is also higher on the probability scale, i.e., we are more likely to come across cases of NC in coordinate contexts than it is the case with non-coordinate contexts. This difference, however, has no bearing on the rate of change which remains unaffected throughout all the six stages. However, visual inspection of Figures 3 and 4 indicates that there are some inadequacies at the end points on the graph standing mainly for stages 5 and 6 in our categorisation.

The Genmod Procedure

<table>
<thead>
<tr>
<th>LR Statistics For Type 1 Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
</tr>
<tr>
<td>Intercept</td>
</tr>
<tr>
<td>stage</td>
</tr>
<tr>
<td>function</td>
</tr>
<tr>
<td>function*type</td>
</tr>
</tbody>
</table>

Table 7: The GENMOD procedure.

Figure 3: The plot of data for non-coordinate contexts with the fitted logistic regression line superimposed.
6.2 Micro contexts: GFs

We shall now analyse and compare:

i. The rate of decline of NC in objects in non-coordinate contexts to the one in coordinate contexts.

ii. The rate of decline of NC in adjuncts in non-coordinate contexts to the one in coordinate contexts.

iii. The rate of decline of NC in objects to the one in adjuncts in non-coordinate contexts.

iv. The rate of decline of NC in objects to the one in adjuncts in coordinate contexts.

An analysis of the same process in micro contexts, GFs, as they occur within macro contexts indicates that the decline of NC takes place at the same rate. The logistic regression lines (see Appendix C for graphs) obtained are once again parallel, meaning that the frequency of n-words in Objects as in Adjuncts was declining at the same rate in non-coordinate constructions and in coordinate constructions. However, once again there is a lack of fit observed on these graphs at stages 5 and 6.

To summarise, our data mainly adheres to the claims made about the Constant Rate Effect across contexts. However, the model as pointed out shows some lack of fit at different stages in different contexts. Accordingly, a plot of residuals and predicted values (i.e. the difference between original values and predicted values) was applied to this model. This plot is normally applied to test the validity of a certain model. Having applied this plot, it clearly indicates some quadratic effect, i.e. evidence of curvature. This raised the issue of considering another logistic model in order to manipulate any lack of fit and the curvature that might exist. Accordingly, another statistical model was applied to analyse the same data and compare it with the previous one. The following section provides a detailed analysis of our data based on another logistic model.

7 An alternative analysis

After fitting a model to a set of data, it is natural to enquire about the extent to which the fitted values of the response variable under the model compare with the observed values. This aspect of the adequacy of a model is widely referred to as goodness of fit (cf. Collett 2003). As suggested above, the logistic linear model as applied to our data shows a lack of fit, mainly at the last two stages. Visual inspection of the graphs indicate that the fit for the first four stages seems to be good which is not the case with the last two stages where there is often some sort of deviation of linearity. It is not clear why there is a lack of fit in the last two stages.
From the plots where linear trends are assumed there seems to be evidence of curvature, which was modelled empirically by including a stage square term ($x^2$) in addition to the 'stage' linear term. This essentially corresponds to modelling a quadratic 'stage' effect. If the quadratic effect for stage obtains, the lines on the graphs can take any other shape apart from straight lines and thus the change can no longer be said to be constant across time. Thus, a slightly more complicated model is applied, one that is in close agreement with the observed data at extreme values. A model for the relationship between the true observed probabilities, $p$, and an explanatory variable, 'stage' in our case, is called a polynomial logistic regression model. The general form of the model is given in (17).\(^\text{13}\)

(17)

\[
\logit = \log \left( \frac{x}{1-x} \right) = \beta_0 + \beta_1 \text{type} + \beta_2 \text{stage}
\]

The applied quadratic model as applied to our data yielded the following results which maintain the same conclusions in terms of the effect of 'type', i.e. our two grammatical constructions, and the effect of 'function', i.e. our two grammatical functions and their interaction with the stage variable. However, the results in terms of the effect of stage show a highly significant chi-square value (<.0001, in 18 below). This highly significant figure makes evident the effect of curvature and simply rules out the validity of the previously used model in analysing our data. These results are displayed in Table 8.

<table>
<thead>
<tr>
<th>LR Statistics For Type 3 Analysis</th>
<th>source</th>
<th>Num DF</th>
<th>Den DF</th>
<th>F Value</th>
<th>Pr &gt; F</th>
<th>Chi Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>stage</td>
<td>1</td>
<td>19</td>
<td>3.86</td>
<td>0.0644</td>
<td>3.86</td>
<td>0.0496</td>
<td></td>
</tr>
<tr>
<td>type</td>
<td>1</td>
<td>19</td>
<td>14.43</td>
<td>0.0012</td>
<td>14.43</td>
<td>0.0001</td>
<td></td>
</tr>
<tr>
<td>stage*stage</td>
<td>1</td>
<td>19</td>
<td>22.99</td>
<td>0.0001</td>
<td>22.99</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>function</td>
<td>1</td>
<td>19</td>
<td>0.57</td>
<td>0.4592</td>
<td>0.57</td>
<td>0.4500</td>
<td></td>
</tr>
</tbody>
</table>

Table 8: LR statistics for a Type 3 analysis.

When applied to our data, this model reveals, based on the statistical analysis presented above, that we have more significant values than the ones yielded by the linear model previously applied. What is crucial here is the quadratic term used to check how linear the trend is in a given observed change. According to the new non-linear model, this particular effect is highly significant as the values in bold (Table 8) indicate. This indicates that the observed probabilities of use of n-words in different contexts are well fitted by the quadratic logistic regression model. Although visual inspection of some of the graphs indicates that there is still some lack of fit in some middle stages, the model is still statistically deemed to be more adequate than the previous one.

\(^{13}\)p stands for the true probabilities of n-words.
Figure 7: The plot of data for adjuncts (Function 2) in non-coordinate and coordinate contexts.

Figure 8: The plot of data for objects and adjuncts in non-coordinate contexts (Type 1).

Figure 9: The plot of data for objects and adjuncts in coordinate contexts (Type 2).
Note that in spite of the fact that we now have curves instead of straight parallel lines standing for a linear trend, the same rates of decline of NC are still maintained in all contexts, i.e. the same behaviour is observed in the different considered contexts. Figure 5 shows that the rate of decline of NC in macro contexts is the same for both non-coordinate and coordinate contexts, while it indicates that it is not constant across time. The rate of decline of NC in both Objects (Figure 6) in both grammatical constructions is the same, and the same is true of Adjuncts (Figure 7) in both construction types. A comparison of the rate of decline of NC in both Objects and Adjuncts as they occur within each grammatical construction (Figures 8 and 9) separately also yields highly significant results indicating that the two GFs behave in exactly the same way within each construction type throughout all our stages. Data show that for a fixed level of stage, the rate of change is the same in all contexts in some sort of parallel curvature.

To summarise, we have seen that the linear model as applied to our data shows some lack of fit at stages 5 and 6 and inadequacies (cf. Figures 3 and 4) are evident from visual inspection of the graphs. Accordingly, another logistic model is used in order to manipulate the observed lack of fit. The new non-linear model still shows some inadequacies in some stages. However, the model is statistically shown to provide a better fit for our data (see 19 above). This then makes our findings only partly compatible with Kroch’s (1989, 1994) CRH. A full discussion of the implications of this alternative model for the CRH is provided in the following section.

8 The Context Constancy Principle

The extent to which the CRH can account for the observed linguistic change will be questioned in the light of the outcome of the two statistical models as applied to our data. We would like to argue that the CRH does not provide an adequate model in order to account for the observed decline and loss of NC in terms of the claim made concerning the constant rate effect across time. Kroch (1989, 1994) argues that changes proceed at the same rate in all contexts, contra Bailey who claims that the change proceeds faster in more favourable contexts and suggests that linguistic changes not only proceed at the same rate across contexts but also across time. This constancy effect across time would statistically be represented as straight (for time effect) parallel (for contexts effect) lines. However, our data can be better modelled by curvature. On modeling this curvature, more significant chi-square values were obtained and the model clearly indicates a better fit for our data. Although the figures from the second model now illustrate curves, rather than straight lines, the findings uphold Kroch’s constancy effect across contexts. The new model and the logistic curves indicate the same pattern of decline of NC in all contexts; this is demonstrated by the parallel curves obtained, something we shall refer to as parallel curvature. This is in accordance with the claims made by Kroch (1989, 1994) that linguistic changes proceed at the same rate in all contexts but contradicts his claim that the change will proceed at the same rate across time in all contexts because none of the contexts we have considered during the process of decline and ultimate loss of NC shows a constant rate effect across time, i.e. we do not have the same rate of decline of NC, say, between 1450 and 1499 and between 1500 and 1550. Accordingly, we would like to argue that what constitutes a principle of linguistic change is not constancy across time but constancy across contexts. Based on this distinction, we shift focus to contexts similar behaviour as the key issue in linguistic changes. We shall call this context constancy principle (CCP). According to this principle, a context constancy effect (CCE) should be obtained when studying the rate of change across a set of linguistic contexts. This principle adheres to Kroch’s Constant Rate Hypothesis in some of its aspects but not in others. Our data show that the Constant Rate Effect was not attainable throughout all the observed stages, and thus cannot adequately account for some aspects of the change. The decline of NC did not follow a constant rate across time in any of the macro and micro contexts. What is, however, consistent with the CRH is the fact that the same pattern of change is obtained for all contexts. This is a crucial issue as it provides an explanation to the fact that these contextual surface changes are triggered by a change in a single underlying parameter. These findings then indicate that whether the change is constant across time (parallel lines) or non-constant across time (curves) is not crucial to the linguistic theory; what is crucial, however, is whether the change is constant across contexts as this bears significant implications for the linguistic theory in general.

To summarise, Late Middle and Early Modern English speakers were endowed with two grammars, i.e. they had an old grammar which is [+NC] and a new one which is [-NC], and accordingly varied their use and alternated between both systems. These two grammars coexisted and were competing for the same structural position. In the case of our study, two forms which have identical meaning and the same discourse function were alternating. The changes in frequency of use of n-words and negative polarity items that we have established over a period of one and a half centuries indicate a shift in language users’ overall tendency to use one grammatical option over another in their language output. This changing tendency is reflected in the changes in surface contexts where usage frequencies can be measured. The unity of the change, however, Kroch (1994) argues, is defined at the level of the grammar not at the level of surface contexts. This study supports the idea put forward by Kroch (1994) that syntactic change, the outcome of a diachronically unstable alternation, proceeds via competition between incompatible grammar options which are used interchangeably.
9 Lexical ambiguity and parameter change

The empirical record shows that the use of NC underwent some gradual statistical shifts through the Late Middle and Early Modern English period. If the consulted texts are reliable indicators, in the Late Middle English NC was almost categorically used across contexts, but at the end of the sixteenth century, variation stopped to exist when the [±NC] grammar became obviously more favoured than the [+NC] variety which gradually became less frequently attested in the texts through the Early Modern English period. The situation became crucially ambiguous when two readings became available and ambiguity arises between: a parameter setting that allows for (1) a clause with a single negative meaning, and (2) a double negation. With n-words being lexically re-analysed as purely [-neg] it became impossible to obtain the intended reading of a single negation (late sixteenth century). The fact that this doubly negated reading is available explains the development into a single option with a new parameter setting disallowing the co-occurrence of two or more n-indefinites.

Before the parameter change, n-words were ambiguous between NPIs, elements that are dependent on other elements higher in the clause for their interpretation, and negative quantifiers (NQs). With these utterances becoming less frequent, the parameter allowing for n-words to co-occur with other negative elements higher in the clause was reset and n-words came to be re-interpreted as NQs after behaving as NPIs. They became independent elements in the sentence and can convey negation meaning without 'seeking help' from another negative element, either a sentential negator or another n-word. This lexical reanalysis in the nature of n-items only took place when the frequencies of n-items were showing a constant decline and evidence for their introduction in negative contexts were not robust enough to keep the same parameter-setting (the second half of the sixteenth century). With the parameter for n-words being reset, n-words stopped being ambiguous between NPIs and NQs. The lexical reanalysis and the resulting new negative quantifiers status that n-words have acquired enabled them to express the negative meaning without having to be c-commanded by a second negative element in the clause.

This change took place across the board, even in contexts where NC took the pattern of two or more n-words co-occurring in the same sentence. In other words, n-items declined dramatically in contexts where they had to be licensed by another n-item higher in the clause and were replaced by NPIs in the contexts considered. Our data indicate that the lexical reanalysis of n-items was across the board, i.e. in all the syntactic contexts considered. This is in accordance with Haspelmath (1997: 219) who argues that, if in a language negative pronouns do not co-occur with verbal negation, they also do not co-occur with each other, and conversely.

To summarise, we suggest that the use of n-words in structural positions where only NPIs are found today was only a stage in the transition that the system of negation was experiencing in Middle and Early Modern English. During this transitional stage, the role of n-words was to reinforce the negative meaning after the loss of bipartite negation, namely the use of ne...not. This caused ambiguity to arise between a single and a double interpretation of the negative sense. What happened is that n-items, which became ambiguous between an NPI and an NQ interpretation, underwent a lexical reanalysis and have now acquired a new single status, that of NQs, as from the time they could stand on their own and express the negative meaning. The change in the status of n-items is a change that affected a whole set of lexical items and is better treated as a case of parameter resetting as it allows us to account for a variety of phenomena economically. We have shown how certain surface changes can be accounted for in terms of the resetting of a single parameter, that of the features attributed to n-indefinites; a parameter which applied across the board.

10 Conclusion

Our findings indicate that the change is natural and adheres to principles of linguistic change. This study shows that the decline of n-words in all the observed contexts is consistent with Bailey’s (1973) first principle of language change, stating that linguistic changes tend to follow an S-shaped curve which emphasizes the slow-fast-slow pattern in the process of the change. This pattern of change was observed in both macro contexts (grammatical constructions) and micro contexts (GFs), where S-curves were obtained.

These findings also clearly show that the process of decline of NC followed a natural path and that it is triggered by theory-internal motivations. The change followed the same pattern across contexts. All macro and micro contexts behaved similarly in a PARALLEL CURVATURE. However, a constant rate effect across time, based on Kroch’s (1989, 1994) CRH, was not obtained. Consequently, an alternative principle of linguistic change, which enables us to account for the loss of NC, is suggested. The CONTEXT CONSTANCY PRINCIPLE shifts focus to contexts’ similar behavior and emphasizes the pattern of change in different contexts, which is more crucial to the theoretical accounts of language change as it reveals issues related to the nature of the change and its driving forces. Constancy across time, this study shows, is not a requirement; it may or it may not be obtained but this does not change the fact that contexts behave similarly, which has crucial implications for the linguistic theory in general terms.

The fact that the decline of NC followed the same pattern, i.e. took place at the same rate across the board suggests that the change is internally-driven. The period of variation that preceded the change was manifested by the competition between mutually exclusive options, namely n-words and negative polarity items in negative contexts. This variation culminated in the establishment of NPIs in these contexts. This study argues based on our data analyses that this competition resulting in the loss of n-words in
negative contexts is the outcome of a lexical reanalysis that n-words underwent. The polysemous status of N-ITEMS between NPIs (originally) and NQs (at a later stage) led to their re-interpretation. The disambiguation of negative contexts where negation could be single or double only took place when N-ITEMS were reanalysed as NQs. This lexical reanalysis is better seen in terms of a lexical parameter resetting, i.e. a formal change affecting a whole class of lexical items. First, it involves the acquisition of a new grammatical feature [+ Neg] by n-words. Second, it allows us to account for a variety of observed surface phenomena in terms of this new parameter setting, namely the lexical reanalysis of n-words from a status whereby they used to behave as negative polarity items and thus rely on other negative elements to license them, to a status where they behave as negative quantifiers, purely negative syntactic elements.

Appendix A

Primary Sources


Appendix B

This Appendix contains a list of the texts within each stage.

Stage 1: (1450-1474)
- The Paston Letters (1425-1495)
- The Plumpton Letters (1433-1551/2)
- The Stonor Letters (1290-1483)

Stage 2: (1475-1499)
- The Cely Letters (1472-1488)
- The Paston Letters (1425-1495)
- The Plumpton Letters (1433-1551/2)
- Christ Church Letters (1334-1520)
- The Stonor Letters (1290-1483)

Stage 3: (1500-1524)
- The Plumpton Letters (1433-1551/2)
- Letters of Royal and Illustrious Ladies of GB (1451-1550)
- Christ Church Letters (1334-1520)
- The Clifford Letters I (1510-1549)
- Letters of Richard Fox (1486-1527)

Stage 4: (1525-1549)
- The Plumpton Letters (1433-1551/2)
- The Clifford Letters II (1525-1547)
- The Clifford Letters I (1510-1540)
- The Lisle Letters (1533-40)
- Letters of Royal and Illustrious Ladies of GB (1451-1550)
Appendix C

Figure 10: The plot of data for Objects in non-coordinate contexts with the fitted logistic regression line superimposed.

Figure 11: The plot of data for Objects in coordinate contexts with the fitted logistic regression line superimposed.
References


