The loss of negative concord in Standard English: Internal factors

AMEL KALLEL
Reading University

ABSTRACT

This study readdresses the loss of Negative Concord (NC) in Standard English. A detailed study of negation in Late Middle English and Early Modern English reveals that the loss of NC was a case of a natural change triggered by some internal factors. A close study of *n*-words in negative contexts and their ultimate replacement with negative polarity items (NPIs) 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 (see Contra Kroch's Constant Rate Hypothesis [CRH], 1989). Context behavior 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.

Although Modern Standard English (1800–1920) is not characterized by the operation of negative concord (NC), in some dialects of English[1] and in certain older forms of the language the operation of negative concord is much stronger. Modern Standard English exhibits a uniform [−NC] system[2], whereas earlier forms of English are characterized by the phenomenon of NC. References to this effect may be found in most general studies of Middle English (1100–1500), as well as in those of Early Modern English (1500–1800) (Barber, 1997:283; Burnley, 1983:61). These periods exhibited variable use of [+NC] and [−NC] systems. The following two examples illustrate the (−NC) and (+NC) systems, respectively:

(1) He did*n't* hurt *any*body
(2) John did*n't* hurt *no*body

It has been largely assumed that the loss of NC was the outcome of prescriptive views on language use (cf. Cheshire 1982:63), and of taking Latin, a [−NC] language, as a model for English grammar. Because of these assumptions, the issue of *why* NC was lost in Modern Standard English was not given enough

attention. This issue will be readdressed 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 Early Modern English period.

NEGATION IN PREMODERN 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, hence the name NC which accordingly excludes concord cases where "any" items are used together with another negative element. Consider the following examples:

(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)

By the Late Middle and Early Modern English periods, which in this study stand for the period running from 1450 to 1600, speakers had an alternative option, which now makes use of *any*-words in contexts where *n*-words were used, and competition between these two variants, *n*-words and *any*-words, arose. There was a period of variation wherein both grammars, [+NC] and [−NC], coexisted. The notion of coexisting grammars, which implies that the rate of change in different surface contexts reflecting a single underlying parameter change is the same, is supported by evidence in the literature. This is known as Kroch's (1989) 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 CRH.

THE CONSTANT RATE HYPOTHESIS

Kroch (1994) argued 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, that is, 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. The Constant Rate Hypothesis, which is a statement of a statistical null hypothesis, states that different environments involved in a change show the same (logistic) rate of change over time, that is, the change has the same regression slope in all contexts. This means that there is no interaction between time and context, hence the null hypothesis. According to Kroch (1994), the grammatical analysis that defines the contexts of a change is quite abstract. He suggested 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, that is, they have the same slope S, which stands for the rate of change.

Bailey's (1973:77) first principle of linguistic change 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.

DATA SOURCES AND CATEGORIZATION

This study makes use of a corpus of Late Middle English and Early Modern English texts created for the purpose of this study (discussed later). The texts selected range in date from the second half of the 15th century until the end of the 16th century, that is, from 1450 up to 1599. Only private correspondence texts were used as a source of data because of the nature of the change we are addressing. Private letters are thought to be closer to the vernacular, from which changes are more likely to spring, than other genres. Other written documents belonging to that period tend to lack in speech-like features and show elaborate structure and at times a high degree of literary style. Accordingly, all private correspondence belonging to 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. Translated letters and letters that have a quasi-public nature, namely, those by and to members of the Royal Family, were excluded from our analysis. The following is a list of all the sources consulted, categorized by period.

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 (1510–1549)
The Lisle Letters (1533–1540)
Letters of Royal and Illustrious Ladies of GB (1451–1550)
The Letters of Thomas Greene (1528–36/1536–38)
The Willoughby Letters (1525–1549)
Letters of Thomas Cromwell (1533–1540)

Stage 5: (1550–1574)
The Letters of Thomas Bentham (1560–1561)
The Letter-Book of John Parkhurst (1571–1575)
The Letters of Richard Scudamore (1549–1555)
The Papers of Nathaniel Bacon of Stiffkey (1556–1577)

Stage 6: (1575–1599)
The Leycester Correspondence (1585–1586)
Two Elizabethan Women (1575–1611)
The Hutton Correspondence (1565–1600)
The Letters of Lady Dorothy Bacon (1597–1622)
Letters of John Holles (1587–1637)
Hastings Letters (1574–1609)
Gossip From a Muniment Room (1574–1618)

It is a well-known fact that changes are likely to be implemented in speech before they spread to the written form of a language. It has also been suggested that a study dealing with language change based on written texts alone would give an unsatisfactory and incomplete picture of the change. Negation is undeniably a field rich in its variety of expression. Practically, the only way to get information on the expressions typical of the spoken language of past centuries is through observations based on the types of writing that can be assumed to be least distant from the oral mode of expression.

Different genres were available for this study. As in Present-day English, Early Modern English reflects different styles, ranging from the most formal to the least formal ones. There are some written documents belonging to that period that are thought to be closer to the vernacular than other genres, those relating to activities such as trials, drama, private correspondence, and diaries. There are also other written documents, but these are literary in nature and at times very limited. This genre lacks speech-like features and its structure is much more elaborate than spontaneous speech.

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. Originally, we looked at negative, interrogative, and conditional clauses, but the last two structures were low in absolute frequency and accordingly were excluded from this study. These cases were categorized in terms of grammatical constructions, namely, noncoordinate (examples 5–6) and coordinate negative environments (examples 7–8).

(5) 'that ye wryt *not* to me *no* letters' (*The Cely Letters*: 10)
(6) 'that I shulde *not* take *any* writt ageynst theym' (*Christ Church Letters*: 66)
(7) 'I had none, *ner* he delyueryd me *none*' (*The Cely Letters*: 18)
(8) '*ne* bounde for *any* tryell of your seid' (*Christ Church Letters*: 41)

They were also categorized in terms of two grammatical functions, namely, Objects and Adjuncts, as they occur within these two grammatical constructions. The first two examples illustrate Objects as they occur in both noncoordinate and coordinate contexts. Examples 11–12 illustrate Adjuncts as they occur in both of these constructions.

 (9) 'that ye wryt *not* to me *no* letters' (*The Cely Letters*: 10)
(10) '*Nor* send me *no* ster.[3] money' (*The Cely Letters*: 12)
(11) 'woll *nott* dyshease yow off yowre howsse *no* lenger' (*The Cely Letters*: 170)
(12) '*ne* forthere prosede in *no* seche matere' (*The Paston Letters*: 82 & 137)

In other words, all instances of *n*-items in positions where they c-commanded[4] all instances of *n*-items and corresponding NPIs occurring in object and adverbial grammatical functions within noncoordinate 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 summarized as follows:

(13) NEG[5] + *n*-item/*any*-item
(14) *N*-item[6] + *n*-item/*any*-item

This study quantifies the occurrences of NC as the structural unit. Differences are quantitatively measured, and our six stages,[7] 25 years each, will be distinguished by the frequency with which some variant occurs in one stage as opposed to another. Differences in any one variable are tested against the background of their frequencies rather than in terms of their absence or presence. A crucial step in our analysis is to discover the rates of change in the use of NC in both noncoordinate and coordinate contexts and to compare them. According to the CRH (Kroch, 1989), 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. Following these claims, we would then expect to find the same rate of change in these two grammatical constructions. We used the Proc Genmod

procedure within the SAS statistical package to fit data into the logistic regression (Collett, 2003). The model must adequately fit[8] the observed probabilities for the fitted linear logistic model in order for it to be satisfactory. By fitting empirical data (i.e., percentages), to the logistic function, we can determine whether the rates of change in different contexts are the same or different.

THE LOGISTIC REGRESSIONS

This study addresses the question of whether the change as observed in different contexts follows an *S*-shaped curve (cf. Bailey, 1973; Kroch, 1989). Labov (1994:65) showed how an *S*-curve is produced by the cumulative frequencies of the binomial distribution. The logit (i.e., the logistic transform shown in [15]) 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.

(15)

$$\text{Logit} = 1n\ \frac{p}{1-p} = k + st.^{9}$$

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, such as 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.

DATA ANALYSIS

In this section we analyze the shift in the percentages of use of NC in noncoordinate 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[10] until NC virtually disappeared at stage 5, with 3.1%, and went completely out of use at stage 6, the last quarter of the 16th century, with only 0.6% usage. This decline in the use of NC in noncoordinate contexts meant a corresponding gradual increase in the use of negative polarity items in contexts in which *n*-words once were used. In stage 1, negative polarity items were used only 16.6% 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 noncoordinate constructions follows an *S*-shaped curve. Table 2 summarizes the frequencies of use of NC and NPIs in coordinate contexts throughout the six

TABLE 1. *The frequency of* n-*words and NPIs in noncoordinate constructions by stage*

Stage	Type One Noncoordinate		
	NC	NPIs	Total
Stage 1	105	21	126
	83.3%	16.7%	
Stage 2	91	25	116
	78.4%	21.6%	
Stage 3	13	14	27
	48.1%	51.9%	
Stage 4	81	106	187
	43.3%	56.7%	
Stage 5	5	155	160
	3.1%	96.9%	
Stage 6	1	156	157
	0.6%	99.4%	

Noncoordinate constructions will be referred to as type 1 in the graphs; coordinate constructions as type 2.
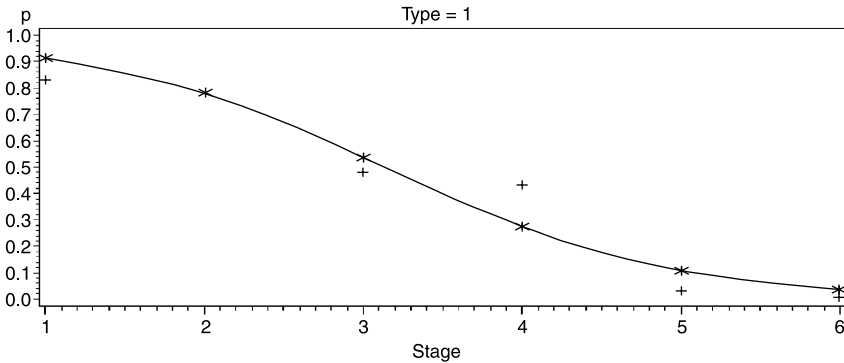

FIGURE 1. The observed data for noncoordinate contexts plotted against the fitted logistic regression.

stages. The picture of the change is very similar to the one observed in noncoordinate 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% in stage 6. The change again reflects the on-going rise in the frequency of negative polarity items in contexts where *n*-words were used. The corresponding percentage in NPIs rose from only 3.6% in stage 1 to 93% in stage 6. Figure 2 indicates that the *S*-shaped curve is also found in coordinate contexts.

TABLE 2. *The frequency of* n-*words and NPIs in coordinate constructions by stage*

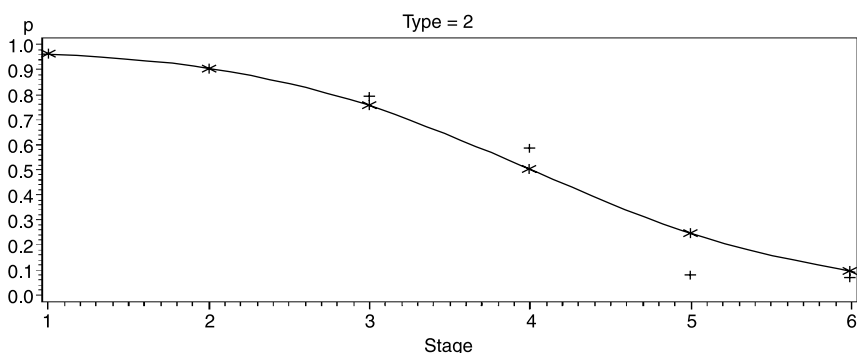|	Type Two Coordinate		
Stage	NC	NPIs	Total
Stage 1	82	3	85
	96.5%	3.5%	
Stage 2	54	6	60
	90%	10%	
Stage 3	23	6	29
	79.3%	20.7%	
Stage 4	64	45	109
	58.7%	41.3%	
Stage 5	4	47	51
	7.8%	92.2%	
Stage 6	3	40	43
	7%	93%	

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 Adjuncts,[11] as they occur within the two grammatical constructions, yields the same results: an *S*-shaped curve. Tables 3 and 4 record the occurrences and percentages of both NC and NPIs in negative clauses as Objects in noncoordinate 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. Cases of NC as Objects occurred with a frequency of 80% in noncoordinate contexts and 95% in coordinate contexts at stage 1 and then dropped to 1.3% and 3.6% at stage 6 in both grammatical constructions, respectively.

TABLE 3. *The frequency of* n-*words and NPIs in Objects in noncoordinate contexts by stage*

	Object Noncoordinate		
Stage	NC	NPIs	Total
Stage 1	44	11	55
	80%	20%	
Stage 2	49	13	62
	79%	21%	
Stage 3	5	7	12
	41.7%	58.3%	
Stage 4	27	49	76
	35.5%	64.5%	
Stage 5	3	79	82
	3.7%	96.3%	
Stage 6	1	74	75
	1.3%	98.7%	

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

	Object Coordinate		
Stage	NC	NPIs	Total
Stage 1	38	2	40
	95%	5%	
Stage 2	30	4	34
	88.2%	11.8%	
Stage 3	14	2	16
	87.5%	12.5%	
Stage 4	36	20	56
	64.3%	35.7%	
Stage 5	2	28	30
	6.7%	93.3%	
Stage 6	1	27	28
	3.6%	96.4%	

The change in NC in Adjuncts (see Tables 5 and 6) follows the same pattern. We notice an on-going 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 apply the Constant Rate Hypothesis and find out whether the decline of NC occurred at the same rate in all studied contexts, as well as 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

TABLE 5. *The frequency of* n-*words and NPIs in Adjuncts in noncoordinate contexts by stage*

	Adjunct Noncoordinate		
Stage	NC	NPIs	Total
Stage 1	59	9	68
	86.8%	13.2%	
Stage 2	40	12	52
	77%	23%	
Stage 3	7	7	14
	50%	50%	
Stage 4	54	55	109
	49.5%	50.5%	
Stage 5	1	76	77
	1.3%	98.7%	
Stage 6	0	82	82
	0%	100%	

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

	Adjunct Coordinate		
Stage	NC	NPIs	Total
Stage 1	30	1	31
	96.8%	3.2%	
Stage 2	17	2	19
	89.5%	10.5%	
Stage 3	8	4	12
	87.5%	12.5%	
Stage 4	25	21	46
	54.4%	45.6%	
Stage 5	2	14	16
	12.5%	87.5%	
Stage 6	1	13	14
	7.2%	92.8%	

Constant Rate Effect in macro contexts, that is, noncoordinate and coordinate contexts being the larger unit comprising object and adverbial grammatical functions (GFs), which will be analyzed subsequently.

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 earlier 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 in (16).

(16) a.

The GENMOD Procedure						
Source	Num DF	Den DF	*F* Value	Pr > *F*	Chi-Square	Pr > Chi-Square
Stage	1	19	140.69	<.0001	140.69	<.0001
Type	1	19	8.10	0.0104	8.10	0.0044
Function	1	19	0.06	0.8103	0.06	0.8076
Stage * Function	1	19	0.00	0.9756	0.00	*0.9753*

(16) b.

The GENMOD Procedure						
Source	Num DF	Den DF	*F* Value	Pr > *F*	Chi-Square	Pr > Chi-Square
Stage	1	20	125.97	<.0001	125.97	<.0001
Type	1	20	2.12	0.1612	2.12	0.1457
Stage * type	1	20	0.35	0.5607	0.35	*0.5540*

Num DF = numerator degrees of freedom; Den DF = denominator degrees of freedom.

These two tables in (16) provide a statistical analysis of the data in terms of our different independent variables (i.e., stage, type, and function, indicate that stage (time) is an important factor in the observed change) scoring significant chi-square values, which provide grounds for the observed decline of NC throughout our period. They also show that the other factors (i.e., type and function) do not interact with stage, which is indicated by the high chi-square values obtained: 0.975 in the case of stage/function interaction, and 0.554 in the case of stage/type interaction.

Figures 3 and 4 show that the lines are parallel, and accordingly, that the rate of change in both noncoordinate 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, see (17). 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
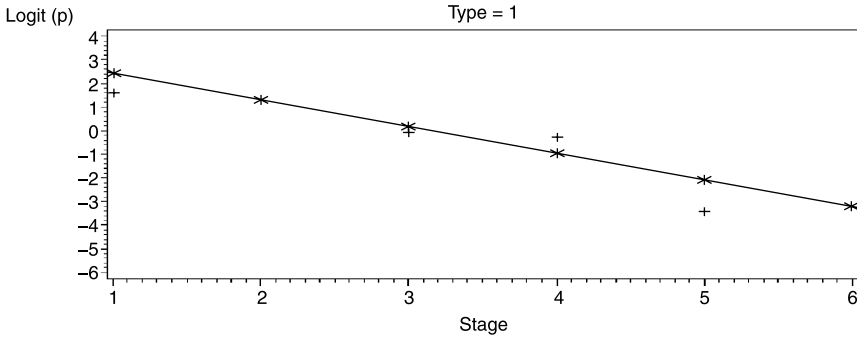
FIGURE 3. The plot of data for noncoordinate contexts with the fitted logistic regression line superimposed.

higher on the probability scale, that is, we are more likely to come across cases of NC in coordinate contexts than in the case of noncoordinate contexts. The difference, however, has no bearing on the rate of change, which remains unaffected throughout all of the six stages. Visual inspection of Figures 3 and 4, however, indicates that there are some inadequacies where the fit of the model does not seem to be satisfactory, mainly for stages 5 and 6 in our categorization.

(17)

				Analysis of Parameter Estimates			
Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	PR > Chi-Square
Intercept	1	4.5154	0.7425	3.0602	5.9706	36.99	<.0001
Stage	1	−1.1264	0.1666	−1.4529	−0.7999	45.71	<.0001
Type 1	1	−0.9790	0.4633	−1.8870	−0.0710	4.47	0.0346
Type 2	0	0.0000	0.0000	0.0000	0.0000		
Scale	0	2.6849	0.0000	2.6849	2.6849		

Micro contexts: Grammatical functions

We shall now analyze and compare:

1. The rate of decline of NC in Objects in noncoordinate contexts to the one in coordinate contexts.
2. The rate of decline of NC in Adjuncts in noncoordinate contexts to the one in coordinate contexts.
3. The rate of decline of NC in Objects to the one in Adjuncts in noncoordinate contexts.
4. 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, grammatical functions, as they occur within macro contexts, indicates that the decline of NC takes place at
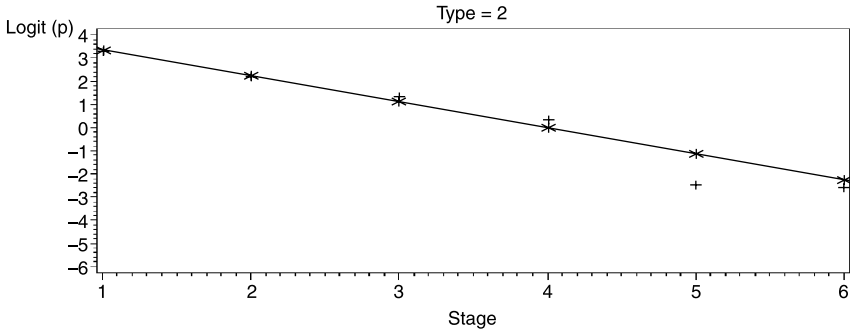
FIGURE 4. The plot of data for coordinate contexts with the fitted logistic regression line superimposed.
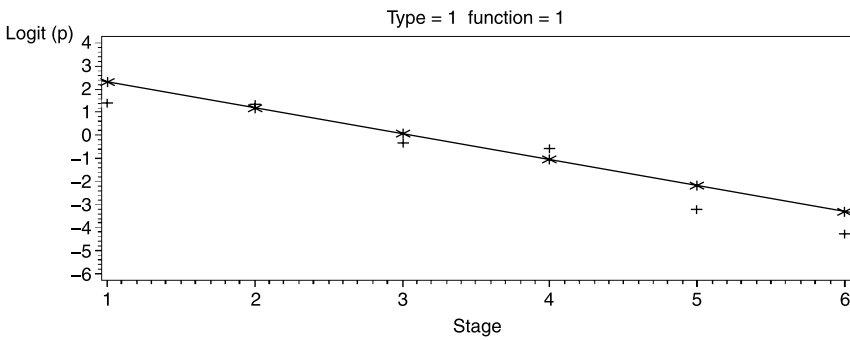

FIGURE 5. The plot of data for Objects in noncoordinate contexts with the fitted logistic regression line superimposed.
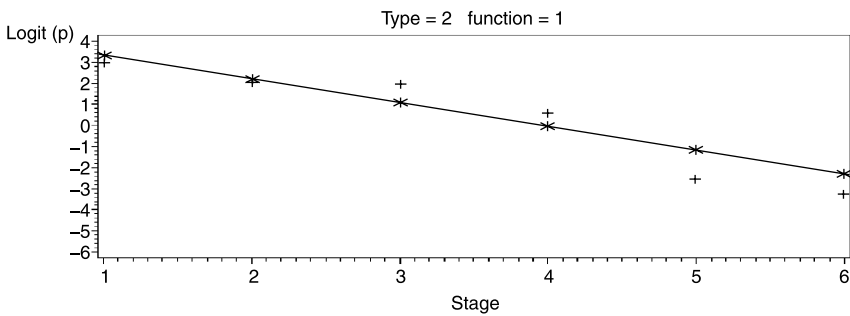

FIGURE 6. The plot of data for Objects in coordinate contexts with the fitted logistic regression line superimposed.
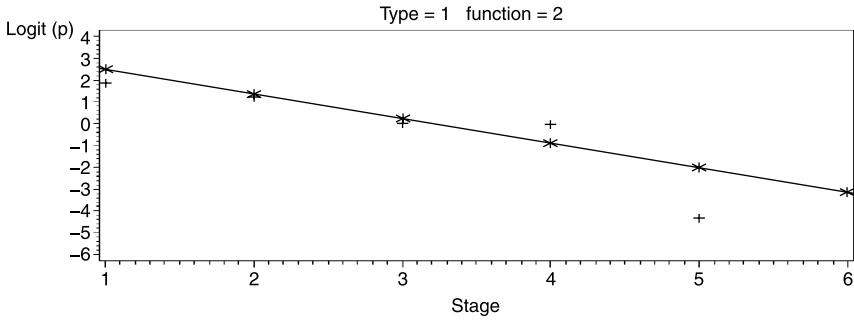
FIGURE 7. The plot of data for Adjuncts in noncoordinate contexts with the fitted logistic regression line superimposed.
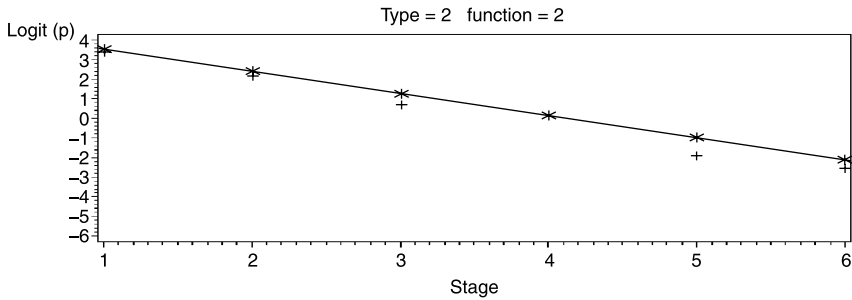

FIGURE 8. The plot of data for Adjuncts in coordinate contexts with the fitted logistic regression line superimposed.

the same rate. The logistic regression lines obtained (see Figures 5, 6, 7, & 8) are once again parallel, meaning that the frequency of *n*-words in Objects, as in Adjuncts, was declining at the same rate in noncoordinate constructions and in coordinate constructions. However, once again there is a lack of fit observed on these graphs at stages 5 and 6.

To summarize, 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, that is, 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 analyze 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.

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 earlier, 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 indicates that the fit for the first four stages seems to be good, which is not the case with the last two stages wherein 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 (18).[12]

(18)

$$\text{logit} = \log_e\left(\frac{\pi}{1 - \pi}\right) = \beta_0 + \beta_1 \text{ type}_i + \beta_2 \text{ stage} + \beta_3 \text{ stage}^2$$

The difference between the first applied model and this one is the quadratic term (x^2) added in this equation (π stands for true probabilities of *n*-words).

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 (19)). This highly significant figure makes evident the effect of curvature and simply rules out the validity of the previously used model in analyzing our data. These results are displayed in (19).

(19)

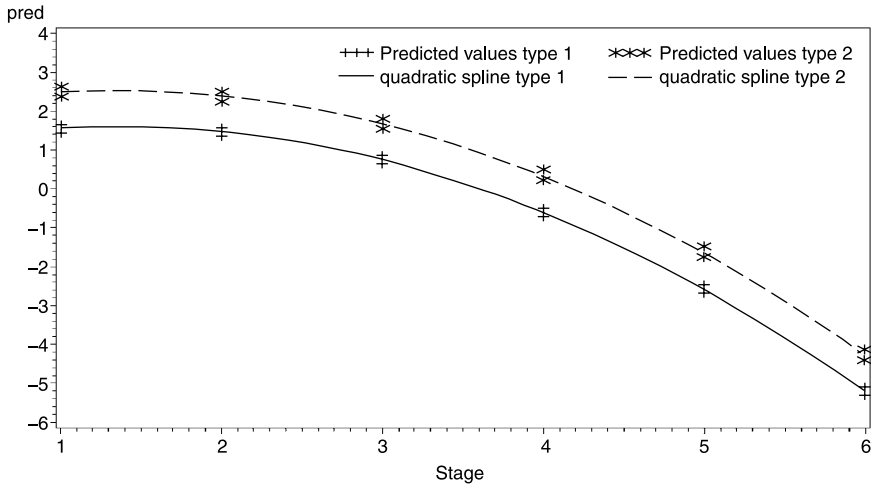| Source	Num DF	Den DF	F Value	Pr > F	Chi-Square	Pr > Chi-Square
Stage	1	19	3.86	0.0644	3.86	0.0496
Type	1	19	14.43	0.0012	14.43	0.0001
Stage * Stage	**1**	**19**	**22.93**	**0.0001**	**22.93**	**<.0001**
Function	1	19	0.57	0.4592	0.57	0.4500

FIGURE 9. The plot of data for noncoordinate and coordinate contexts.

When applied to our data, this model reveals, based on the statistical analysis presented earlier, 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 nonlinear model, this particular effect is highly significant as the values shown in bold type 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. The following figures clearly show that the decline of NC took place at the same rate in all considered contexts.

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, that is, the same behavior is observed in the different considered contexts, which is the null hypothesis. Figure 9 shows that the rate of decline of NC in macro contexts is the same for both noncoordinate and coordinate contexts, while it indicates that it is not constant across time. The rate of decline of NC in both Objects (Figure 10) in both grammatical constructions is the same, and the same is true of Adjuncts (Figure 11) 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 separately (Figures 12 and 13) is also highly significant, indicating that the two grammatical functions 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*.
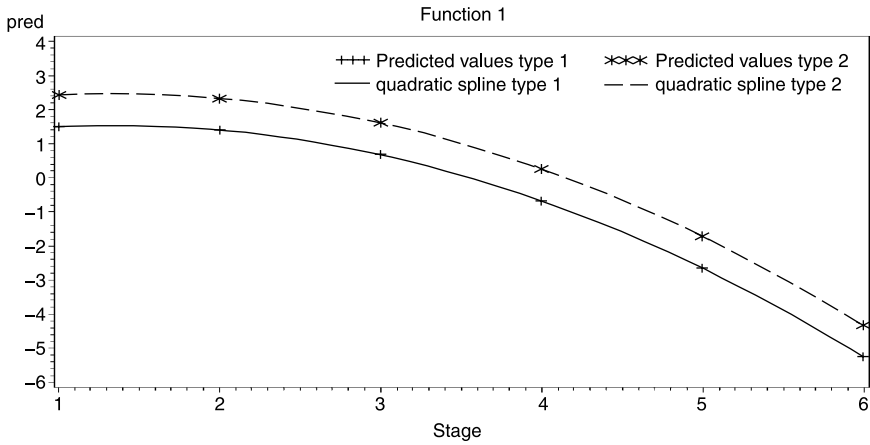
FIGURE 10. The plot of data for Objects (Function 1) in noncoordinate and coordinate contexts.
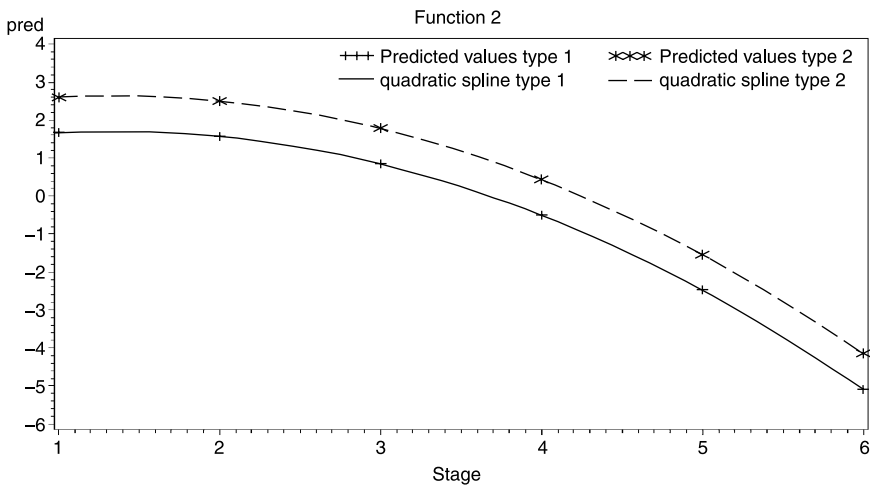

FIGURE 11. The plot of data for Adjuncts (Function 2) in noncoordinate and coordinate contexts.

To summarize, 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 to manipulate the observed lack of fit. The new nonlinear model still shows some inadequacies in some stages. However, the model is statistically shown to provide a better fit for our data, see (19). A full discussion of the implications of this alternative model for the CRH is provided in the following section.
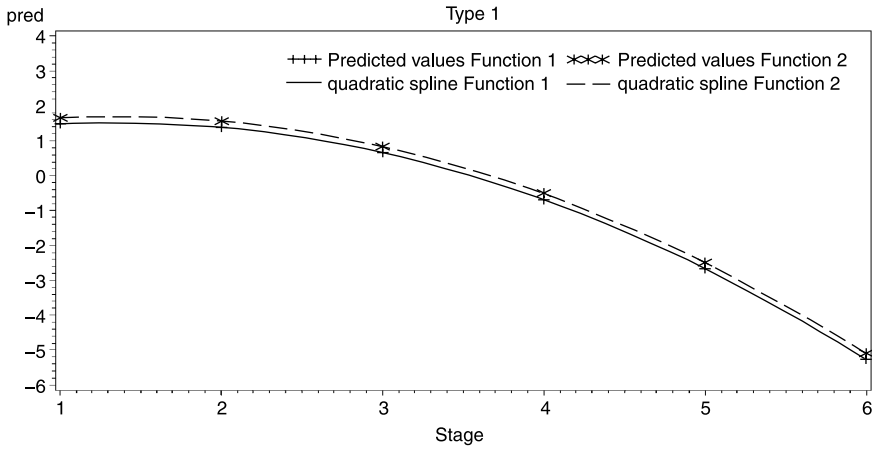
FIGURE 12. The plot of data for Objects and Adjuncts in noncoordinate contexts (Type 1).
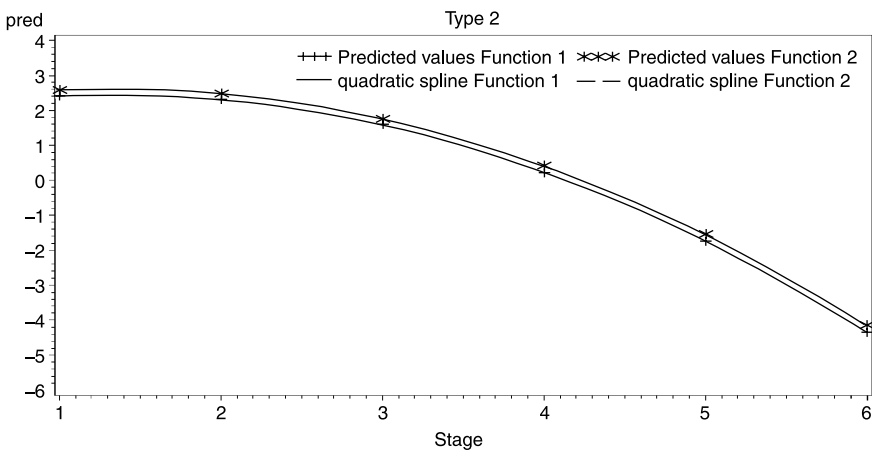

FIGURE 13. The plot of data for Objects and Adjuncts in coordinate contexts (Type 2).

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 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) argued that changes proceed at the same rate in all contexts and suggested that linguistic changes not only proceed at the same rate across contexts, but also across time, which is in fact a statement of a statistical null hypothesis. 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 modelling 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 illus-
trate 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, again,
is a statement of the null hypothesis and is in accordance with the statements
made by Kroch (1989) 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, that is, 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 behavior as the
key issue in linguistic changes. We shall call this the *context constancy princi-
ple* (CCP). According to this principle, a *context constancy effect* (CCE) should
be obtained when studying the rate of change across a set of linguistic con-
texts. This principle adheres to Kroch's Constant Rate Hypothesis in some of
its aspects but not in others. The decline of NC did not follow a constant rate
across time in any of the macro and micro contexts. What is, however, com-
mon to both this study and Kroch's (1989), and is thus consistent with the
CRH, is the fact that the same pattern of change is obtained for all contexts.
The question is: Why is this of any importance? What does the fact that we
now have a constant rate effect across contexts, though not across time, reveal?
This is a crucial issue as it provides an explanation to the fact that these con-
textual surface changes are triggered by a single change, a change in a single
underlying parameter. These findings then indicate that whether the change is
constant across time (parallel lines) or nonconstant 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 linguis-
tic theory in general.

To summarize, Late Middle and Early Modern English speakers were
endowed with two grammars, that is, 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 estab-
lished 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) argued, 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 that are used interchangeably.

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 Late Middle English, NC was almost categorically used across contexts, but at the end of the 16th century, variation ceased to exist when the [−NC] grammar became obviously more favored 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 reanalyzed as purely [+neg] it became impossible to obtain the intended reading of a single negation (late 16th 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 negative 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 reinterpreted as NQs after behaving as NPIs. They became independent elements in the sentence and could convey negative 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 was not robust enough to keep the same parameter-setting (the second half of the 16th 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 argued 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, as well.

To summarize, 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 these *n*-words was to reinforce the negative meaning after the loss of bi-partite 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 that applied across the board.

CONCLUSION

Our findings 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*, showing consistency across contexts but not across time (Kroch, 1989). The suggested *context constancy principle* (CCP) 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 in different contexts, that is, it 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 reinterpretation. The disambiguation of negative contexts where negation could be single or double only took place when *n*-items were reanalyzed as NQs. This lexical reanalysis is better seen in terms of a lexical parameter resetting, that is, 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.

<div align="center">N O T E S</div>

1.   NC has survived in some English dialects, such as the African-American Vernacular English (Labov 1972:785), as the following example illustrates: *Ain't nobody* ever thought about pickin' up *nothin'*.

2.   This study excludes the nonstandard varieties of English that exhibit an NC system.

3.   Ster. = sterling.

4.   A node X c-commands a node Y if:
1.   X does not dominate Y;
2.   Y does not dominate X;
3.   The first branching node Z dominating X dominates Y (Haegeman 1995:4).

5.   NEG refers to the main sentential negator whether it is *not* or any of its earlier forms and spellings, such as *noght, nowght, nat, natte,* etc.

6.   *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 indefinites or negative quantifiers (cf. Zanuttini, 1991).

7.   Stage 1 (1450–1474); Stage 2 (1475–1499); Stage 3 (1500–1524); Stage 4 (1525–1549); Stage 5 (1550–1574); Stage 6 (1575–1599).

8.   The goodness of fit can only be approximated within the statistical model.

9.   The logit of equation (15) produces a straight line, standing for a linear function of time. *s* is the slope of this line; the sharper the slope, the more rapid the change. *k* is the intercept and is related to the frequency of the old/new form at some fixed point in time. The dependent variable only takes values of 0 and 1, but the predicted values for regression take the form of mean proportions or probabilities which are conditional on the values of the independent variables.

   To illustrate the logit transformation, we assume that each case has a probability of having a characteristic and this probability must be estimated. Given this probability, we can then estimate the ratio of probability P to 1-P, i.e. the *odds* of having the characteristic. To proceed using the logit, the probabilities need to be transformed into odds. Probabilities vary between 0 and 1, and express the likelihood of an event as a proportion of both occurrences and non-occurrence. Odds express the likelihood of an occurrence relative to the likelihood of a non-occurrence. Both probabilities and odds have a lower limit of zero, and both express the increasing likelihood of having the characteristic with increasing large positive numbers.

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

11.   Initially, all grammatical functions, such as subjects, embedded subjects, noun post-modifiers, and so forth, were considered, but only Objects and Adjuncts were analyzed statistically, as the others were low in absolute frequency.

12. π stands for true probabilities of *n*-words.

<div align="center">P R I M A R Y   S O U R C E S</div>

Allen, P. S. (ed.). (1929). *Letters of Richard Fox 1486–1527*. Oxford: Clarendon University Press.

Bridgen, S. (ed.). (1990). *The Letters of Richard Scudamore 1549–1555*. Camden Fourth Series 39. The Royal Historical Society. London: Butler & Tanner.

Bruce, J. (ed.). (1844). *Correspondence of Robert Dudley, Earl of Leycester 1585–6*. London: J.B. Nichols & Son.

Cross, C. (ed.). (1969). *The Letters of Sir Francis Hastings 1574–1906*. Somerset Record Society, Vol. LXIX (69). Frome: Butler & Tanner.

Davis, N. (ed.). (1971). *Paston letters and papers of the fifteenth century* (Vol. I). Oxford: Clarendon Press.

Dickens, A. G. (ed.). (1962). *The Clifford Letters of the Sixteenth Century*. Publications of the Surtees Society, Vol. 172. Durham: Andrews and Co.

Hanham, A. (ed.). (1975). *The Cely Letters 1472–1488*. Oxford: Oxford University Press.

Houlbrooke, R. A. (ed.). (1974–1975). *The Letter Book of John Parkhurst, Bishop of Norwich, 1571–5*. Publications of the Norfolk Record Society 43.

Houlbrooke, R. A. (ed.). (1994). *The Letters and Will of Thomas Grene (D. 1545), Rector of Poringland*. Publications of the Norfolk Record Society 56.

Hoyle, R. W. (ed.). (1992). *Letters of the Cliffords 1500–1565*. Camden Fourth Series, Vol. 44. London: Butler & Tanner.

Kingsford, C. L. (ed.). (1919). *The Stonor letters and papers 1290–1483, Volume I*. Camden Third Series, Vol. XXIX (24). London: Offices of the Society.

Kirkby, J. (ed.). (1996). *The Plumpton letters and papers*. Camden Fifth Series, Vol. 8. Cambridge: Cambridge University Press.

Lyell, L. (ed.). (1934). *Miscellaneous Letters: A Medieval Post-Bag*. London: Jonathan Cape.

Newdigate, N. (ed.). N.d. *Gossip from a muniment room: Being passages in the lives of Anne & Mary Fytton 1574–1618*.

O'Day, R. (ed.). (1979). *The Letter-Book of Thomas Bentham, Bishop of Coventry and Lichfield, 1560–1561*. Camden Fourth Series, Vol. 22. London: Butler & Tanner.

Penn, M. A. (ed.). (1965). *Willoughby letters of the first half of the sixteenth century* (Vol. 4). Nottingham: Derry and Sons.

Pickering, W. (ed.). (1843). *The correspondence of Dr. Matthew Hutton*. The Surtees Society. Edingburgh: Laing and Forbes.

Raine. (ed.). (1842). *The Correspondence of Mathew Hutton*. London: J.B. Nichols & Son.

Seddon, P. R. (ed.). (1975). *Letters of John Holles 1587–1637* Thoroton Society. Record Series, Vol. XXXI (31). Nottingham: Derry and Sons.

Sheppard, J. B. (ed.). (1876–1877). *Christ Church Letters 1334–1520*. The Camden Society.

Smith, A. H., Baker, G. M., & Kenny, R. W. (eds.). (1979). *The Papers of Nathaniel Bacon of Stiffkey, Volume I, 1556–1577*, Vol. XLVI (46). Norfolk: Norfolk Record Society.

St. Clare Byrne, M. (ed.). (1981). *The Lisle Letters 1533–1540*, *Volumes I–VI*. Chicago: Chicago University Press.

St. Clare Byrne, M. (ed.). (1981). *Cromwell's Letters incorporated in The Lisle Letters 1533–1540, Volume V*. Chicago: Chicago University Press.

Wall, A.D. (ed.). (1983). *Two Elizabethan women: The Correspondence of Joan and Maria Thynne 1575–1611*. Devizes: Wiltshire Record Society.

Wood, M. A. E. (ed.). (1846). *Letters of royal and illustrious ladies of Great Britain Volumes I–III*. London: Henry Colburn.

REFERENCES

Bailey, C. J. (1973). *Variation and linguistic theory*. Washington, DC: Center for Applied Linguistics.

Barber, C. (eds.). (1997). *Early Modern English*. Edinburgh: Edinburgh University Press.

Burnley, D. (1983). *A guide to Chaucer's English*. London: Macmillan.

Chambers, J. K. (1992). Dialect acquisition. *Language* 68:673–705.

Cheshire, J. (1982). *A sociolinguistic study: Variation in an English dialect*. Cambridge: Cambridge University Press.

Collett, D. (2003). *Modelling binary data* (2nd ed.). London: Chapman & Hall/CRC.

Haegeman, L. (1995). *The Syntax of negation*. Cambridge: Cambridge University Press.

Haspelmath, M. (1997). *Indefinite pronouns*. Oxford: Oxford University Press.

Kroch, A. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199–244.

———— (1994). Morphosyntactic variation. In K. Beals et al. (eds.), *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on variation and linguistic theory*. Chicago: Chicago Linguistics Society.

Labov, W. (1972). *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

———— (1994). *Principles of linguistic change: Vol. I: Internal factors*. Oxford: Blackwell.

Osgood, C., & Sebeok, T. (1954). Psycholinguistics: A survey of theory and research problems. *Journal of Abnormal and Social Psychology* 49:1–203.

Weinreich, U., Labov, W., & Herzog, M. I. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (eds.), *Directions for historical linguistics: A symposium*. Austin: University of Texas Press. 95–189.

Zanuttini, R. (1991). Syntactic properties of sentential negation. A comparative study of Romance languages. Doctoral dissertation, University of Pennsylvania.