University of Pennsylvania Working Papers in Linguistics

Volume 10
Issue 2 *Selected Papers from NWAVE 32*

Article 21

1-1-2005

The Sociolinguistics of DO NOT in the 16th and 17th Century

Anthony Warner

The Sociolinguistics of DO NOT in the 16th and 17th Century

The Sociolinguistics of DO NOT in the 16th and 17th Century

Anthony Warner

1 Introduction

In earlier English there is a long period of competition between the types of (1) and (2). In (1) the negative declarative and the question have their modern form with finite ("supportive" or "periphrastic") DO. In (2) they have the older form with a simple finite verb. See Rissanen (1999) and Denison (1993) for recent general accounts of the history and for bibliography.

(1) a. she does not deserve it, ... (317-4)
    b. why do I spend my time in tittle-tattle with this idle fellow? (215-8)
(2) a. I question not your friendship in the matter, ... (291-23)
    b. Well, madam, how like you it, madam, ha? (301-13)
    (examples from Thomas Otway, *Friendship in Fashion* 1677)

In 1989 Kroch investigated the large database of Ellegård's study of DO (1953), and claimed that the Constant Rate Effect held for the increase in DO up to 1575, that is, that the s-curves of change in different surface contexts were parallel before that date. Subsequently Ellegård's data shows distinct developments in questions and negative declaratives. Kroch took this to be evidence that a grammatical change occurred at this period. I agree with this conclusion, though for another of the reasons he gives, that is, because the decline of unstressed affirmative DO, as in (3), starts at this point.

(3) Why you must know, Frank, having a particular esteem for my family, (the nearest relation of which I would go fifty miles to see hanged) I do think her a very a-- But no more, -- mum, dear heart, mum, I say. (331-35)

What I want to discuss in this paper is the development of negative declaratives in Ellegård's database. This is rather peculiar. This clause type shows a dramatic collapse in the level of DO in the last quarter of the sixteenth century, and an uneven recovery in the seventeenth. So the relative levels of DO in questions and negative declaratives differ sharply at different periods. Before 1575 negative declaratives are 25 years behind affirmative questions in their adoption of DO. After 1600 they lag by over a century.

The transition involves a very rapid decline in the proportion of DO in negative declaratives from 38% in 1550-75 to 24% in 1575-1600 (Ellegård 1953:161ff.). There is also the oddity that another database, the Corpus of Early English Correspondence, shows something different (Nurmi 1999). Here the decline is smaller and it is located rather later, at the beginning of the seventeenth century, not in the last quarter of the sixteenth.

I shall show that both the drop in DO in negative declaratives and its continuing low level is to be explained in large part not as a grammatical but as a sociolinguistic phenomenon. My data is a reconstitution of Ellegård's database of English plays and prose 1500-1710, which I owe to Tony Kroch.[1]

2 Internal Stylistic Differences between Texts

2.1 A Scale of Lexical Complexity

Ellegård's database was not selected with social structuring in mind: he did not pay attention to class or gender in compiling it. His informants are virtually all men who have sufficient education to be literate. It is not therefore possible to treat social variables directly, as it is in the Corpus of Early English Correspondence. But it is possible to look at the internal properties of texts. So I will relate the incidence of DO in the sixteenth and the seventeenth centuries to stylistic level. An adequate measure of this seems to be provided by a combination of the type token ratio and the average word length of Ellegård's texts. Each of these properties separately correlates with the incidence of DO in negative declaratives, and they correlate well with each other. They form a component of Biber's most important dimension of textual variation, indicating the extent to which each text shows "high informational density and exact informational content" (1988:108), to borrow part of his characterization of this dimension. Together they should provide a robust measure of the lexical complexity of texts. So the scales for each were normalized and summed, following Biber's procedure. This gives me a single scale of lexical complexity, which yields intuitively reasonable results.

[1]I want very sincerely to thank Tony Kroch for giving me this database, in an act of straightforward generosity; also Ann Taylor who had compiled most of it, and Celeste Tereszczuk who completed it. I am also grateful to the British Academy who funded a period of research leave which gave me time to investigate this, and to the audience at NWAVE 32 for their comments.

2.2 Do and Lexical Complexity in the Seventeenth Century

When lexical complexity and date are treated as continuous variables in a logistic regression for the seventeenth century, lexical complexity is very highly significant, see (4).[2]

(4) DO in negative declaratives 1600-1710

Factor	Coeff	Prob
Lexical complexity	-ve	$p < 0.0001$

LogLikelihood difference yields chi square of 37.8, df. = 1, Transitivity and date (continuous) also present and significant, N=952, Omits Ellegård's "*know* group" of verbs

The dependent variable here is the choice between DO and finite non-auxiliary verb as in (1) and (2) above.[3] The coefficient is negative, which means that DO in negative declaratives is more frequent in lexically less complex texts, and less frequent in lexically more complex texts. Thus there is less DO in 'higher' registers. For example, Congreve's play of 1693 *The Old Batchelor* has 89% DO in negative declaratives, but his novel *Incognita* (published the preceding year) has 77% DO. Some twenty years earlier Otway's plays have 66% DO, Bunyan's novel, *Pilgrim's Progress* has 36%, and Dyden's late essays have 26% DO: these were all written within a few years of one another.

It is also important to look systematically at the individual texts. Here, it is helpful to set date aside as a variable impacting on the analysis. This can neatly be done by using the rate of change for the period to calculate what the figure for the incidence of DO in negative declaratives for each text would have been in some base year.[4] This calculation has the effect of moving the text along its s-curve of change to the selected year. So Beaumont and Fletcher's play *The Knight of the Burning Pestle* has 64.7% DO in negative declaratives. It was first performed in 1607, and it is possible to work out that the corresponding percentage for 1655 is 81.0%.

[2]The program DataDesk (Velleman 1995) has been used, since it allows for continuous variables.

[3]Other auxiliaries are omitted, as are verbs belonging to Ellegård's '*know* group' which have a very low level of DO. These omissions hold throughout the paper.

[4]The appropriate equation for 1655 is

```
(exp(ln((do/total)/(1-do/total))-(date-1655)*rate)
/(1+exp(ln((do/total)/(1-do/total))-(date-1655)*rate))
```

Essentially this says that if the text had been written under the same conditions at the later date, it would have had 81.0% DO. Similarly, Congreve's play *The Old Bachelor* has 89% DO in 1693; it is on a distinct but very close s-curve, with a corresponding figure for 1655 of 80.4%. When corresponding figures are calculated for all the texts in the period, we can compare the results in respect of stylistic differentiation without interference from date.

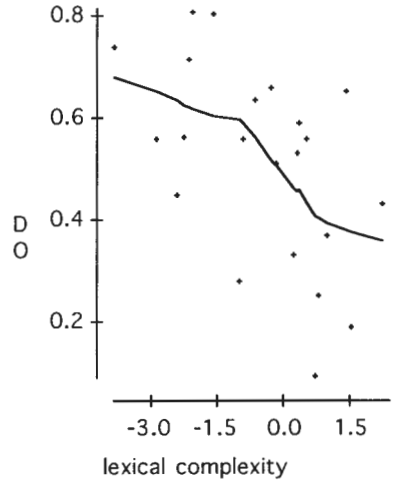

Figure 1: Proportion of DO for 1655 versus lexical complexity, 1600-1710

Figure 1 is a scatterplot for the 22 texts occurring in the period 1600-1710, with proportions of DO estimated for the year 1655. The line is a running average to help you see the trend.[5] You can see that the trend is sharply down from left to right. Three other points are relevant here. First, the linear regression of DO against the scale of lexical complexity is significant ($p < 0.01$). Second, if we divide the graph into four quadrants, at the mid point, the median, of DO% (56%) and of the scale of lexical complexity (-0.2), we

[5]This running average is a 'lowess'. It takes a proportion of the data as a window – here 35% - and produces an average by weighting values more highly as they are closer to the point being established. The running averages given in later figures in this paper are all lowesses. See Velleman 1995 vol 3 chapter 33, Cleveland 1979.

find that most of the texts occurs in two quadrants: low complexity and high DO; high complexity and low DO; see Table 1.

	Texts of low lexical complexity	Texts of high lexical complexity	Total
High DO%	8	3	11
Low DO%	3	8	11
Total	11	11	22
DO% average	61.6%	41.1%	

Table 1: Occurrence of DO in texts of low versus high lexical complexity 1600-1710

Thirdly, we can look at the averages for texts of low and high lexical complexity to get some real idea of the scale of this difference. The average value of DO for the 11 lexically less complex texts is 62%, while for the 11 lexically more complex texts it is 41%. The contrast is striking: on average the lexically less complex texts show 50% more DO than the lexically more complex texts. These three sets of figures for individual texts strikingly confirm the results of the overall logistic regression given in (4), and show that the proportion of DO is higher in lexically less complex texts in Ellegård's seventeenth century database.

2.3 Do and Lexical Complexity in the Period 1500-1575

The situation is, however, interestingly different in the period 1500-1575. Here the results for lexical complexity in a logistic regression including date and transitivity show that this factor is very highly significant ($p < 0.0001$) (see (5)). But now the coefficient is positive: that is, in negative declaratives, there is more DO in texts of higher lexical complexity, less DO in texts of lower lexical complexity. So we have a striking difference between the periods 1500-1575 and 1600-1710: the later period apparently reverses the earlier situation: complex lexis is associated with higher levels of DO in the earlier period, with lower levels of DO in the later period.

(5) Do in negative declaratives 1500-1575

Factor	Coeff	Prob
Lexical complexity	+ve	$p < 0.0001$

LogLikelihood difference yields chi square of 35.3, df. = 1, Transitivity and date (continuous) also present and significant, N=2244, Omits Ellegård's "*know* group" of verbs

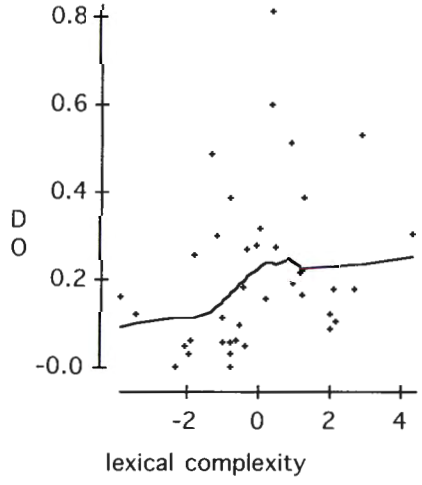

Figure 2: Proportion of DO for 1535 versus lexical complexity, 1500-75

This is also borne out when we look at figures for individuals, setting aside date as before, but this time taking 1535 as a base year. Figure 2 is a scatterplot giving estimates of proportions of DO in negative declaratives for that year, plotted against lexical complexity. There is a clear upward trend from left to right, shown in the running average. Three points to parallel those made before can also be made here. First, the linear regression is significant.[6] Second, if we divide the graph into four quadrants, at the mid point, or median, of each scale, as before, we find that most of the data oc-

[6]Actually, the appropriate significant regression is between the square root of DO% and lexical complexity since this has an approximately normal distribution as DO% itself does not; here $p = 0.027$. But I give the graph for the untransformed variable.

curs in two quadrants: low complexity and low DO%; high complexity and high DO%; see Table 2.

	Texts of low lexical complexity	Texts of high lexical complexity	Total
High DO%	6	14	20
Low DO%	14	5	19
Total	20	19	39
DO% average	14.0%	29.9%	

Table 2: Occurrence of DO in texts of low versus high lexical complexity 1500-1575

Third, comparing the averages is again interesting. The average percentage of DO for the lexically less complex texts is 14%, while for the lexically more complex texts it is 30%: there is twice as much DO in lexically more complex texts. So there is also a considerable difference here. As before we have mutually supportive results from both the logistic regression, and from the figures for individual texts, showing that the incidence of DO in Ellegård's database is higher in texts of higher lexical complexity at this earlier period.

2.4 DO and Lexical Complexity in Negative Declaratives 1500-1710

We can put this together by looking at graphs of change for the whole period 1500-1710, presented in Figures 3 and 4. Here I have divided texts up into two groups of equal size by cutting the scale of lexical complexity at the median. In one group are 39 texts of high lexical complexity; in the other, 38 texts of low lexical complexity. In each case the value for the text is the proportion of DO in negative declarative sentences, and (as before) a running average has been added to help you discriminate the major trends.

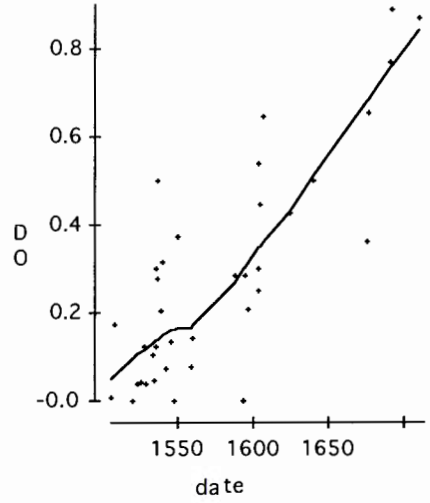

Figure 3: DO versus date in texts of lower lexical complexity

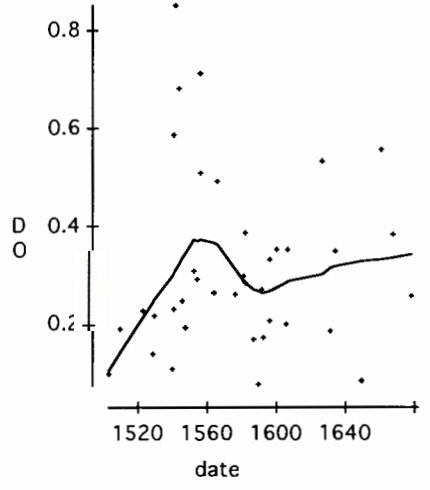

Figure 4: DO versus date in texts of higher lexical complexity

The difference between these graphs is striking. In texts of low lexical complexity there is no sign of the late sixteenth century collapse seen in Ellegård's overall figures, though there is a gap in the data from 1565-85, and an earlier blip. Nor can you see the later irregular and rather flat development his figures show up to the second half of the seventeenth century. But the downward movement in texts of high lexical complexity is dramatic, and a lower level of DO is maintained across the following century. The next graph simply imposes the two running averages on each other: it shows there is a difference in the sixteenth century, and that the situation alters radically as we approach the seventeenth, and remains different.

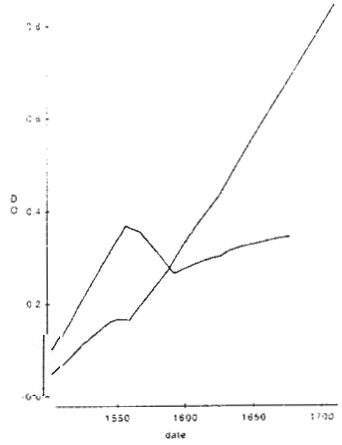

Figure 5: Superimposed running averages from Figures 3 and 4

It is worth noting that this is in line with what Nurmi found in the Helsinki corpus, see Table 3. It is not clear that there is any difference in timing. Note that her categories ("oral" versus "non–oral") were based on a classification by types of genres.

Helsinki 1500-1570	Non-oral > Oral
Helsinki 1570-1640	Non-oral = Oral
Helsinki 1640-1710	Non-oral < Oral

Table 3: Incidence of DO in negative declaratives (Nurmi 1999:147)

It is clear, then, that in Ellegård's database:

1. A higher rate of DO with NOT is found in texts of greater lexical complexity in the period 1500-1575.
2. There is a sharp reversal of this association in the seventeenth century. In this period a lower rate of DO with NOT is associated with texts of greater lexical complexity.
3. The sharp drop in the late sixteenth century in Ellegård's data simply reflects this switch-over which essentially depends on the lexically more complex registers.

In texts of lower lexical complexity there seems to be a pretty steady increase across the whole period. This implies that there is likely to have been a relatively steady underlying development of DO NOT in the vernacular, if we may extrapolate meaningfully from the lexically less complex texts in Ellegård's database. The further drop in 1625-50 which can be seen in Ellegård's figures does not appear in either graph: it is due to sampling differences (as Ellegård himself suggested, 1953:163), since Ellegård's period 1625-50 has a much higher proportion of lexically more complex texts than the periods which precede and follow.

This is all consistent with a change in the relative stylistic values of DO with NOT versus the simple finite with NOT. DO NOT seems to be the more positively evaluated alternative in texts of greater lexical complexity in 1500-75, but to become the more negatively evaluated alternative in seventeenth century texts of greater lexical complexity. This is evidenced, remember, in the written language of men, and we need to bear this restriction in mind. But we might suspect that the fact that the Corpus of Early English Correspondence shows a later decline than Ellegård's database (as reported by Nurmi 1999:148-9, 165ff.) has to do with the different mix of individuals in the two databases, or reflects differences in types of writing: the correspondence corpus contains personal (including private) letters whereas much of Ellegård's corpus was written for publication.

2 Age-grading

Some further evidence which seems supportive of the suggestion that there is a change in evaluation with the seventeenth century can be found in age-grading. Before 1575 the relationship between a writer's age and the incidence of DO in negative declaratives is straightforward. When the date of texts is included in a logistic regression, age is completely nonsignificant, see (6), where the coefficient of age is minimal.

(6) DO in negative declaratives 1500-1575

Factor	Coeff	Prob
Date (of text)	4.6	p<0.0001
Age of author	0.1	n/s p > 0.8

N=1805, Transitivity also present and significant, Omits Ellegårds's "*know* group" of verbs and authors of unknown date of birth, Coefficient of date and age in logit units per century

It looks just like a communal change in Labov's (1994) sense, where individuals change their usage as they grow older, without giving rise to a distinction of apparent time. The older you are, the more you do it, in line with the changing community norm. But in the seventeenth century things are different. Age is a highly significant factor, and its coefficient is negative, see (7): the older you are, the less you do it, although the incidence of DO is continuing to increase.

(7) DO in negative declaratives 1600-1710

Factor	Coeff	Prob
Date (of text)	1.70	p < 0.0001
Age of author	-5.13	p < 0.0001

N=952, Transitivity also present and significant, Omits Ellegårds's "*know* group" of verbs and authors of unknown date of birth, Coefficient of date and age in logit units per century

This too can usefully be examined by looking at the levels of DO for individuals estimated for 1655 in Figure 6. The scatterplot shows a declining trend, clear from the running average. The levelling out depends on one text with a substantially older author (Breton, publishing in 1604 at age 59, who perhaps has not adopted new norms,[7] or who perhaps shows a reversion to the less carefully monitored behaviour patterns of his youth). The linear regression of age against date is significant (p = 0.0076). If the data is split at the medians for age (35.5), and for incidence of DO (56%), the quadrants are as given in Table 4, with a clear contrast, and the average incidence of DO also shows a clear distinction, as you can see in the table.

[7]Breton was born 1545. See Nurmi 1999:173-4 for figures and graphs which imply that the group of individuals born 1520-39 differ from those born 1540-39 and later, in that they did not show a drop in usage in the early seventeenth century, unlike those born later. Breton might have belonged to the earlier group.

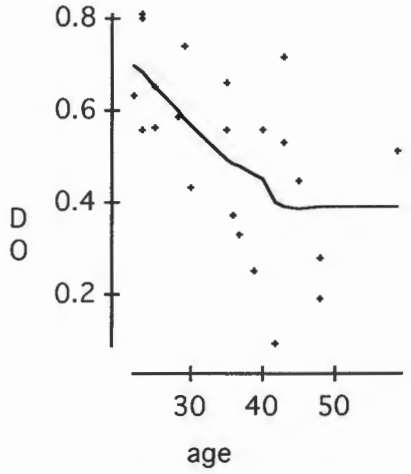

Figure 6: DO for 1655 versus age of author. 1600-1710

	Younger av age 26	Older av age 44	Total
High DO%	10	1	11
Low DO%	1	10	11
Total	11	11	22
DO% average	63.8%	39.0%	

Table 4: Occurrence of DO in texts by younger versus older authors 1600-1710

In his discussion of age-grading, Labov points out that it may be difficult to distinguish from generational change, in which individuals do not change their usage as they age. Is what is going on here generational change? No, for two reasons. The first is that we have apparent communal change for the development of DO in negative declaratives before 1575, since here age is not a significant factor. The same is true of questions, both before 1575 and in the seventeenth century. It would seem very strange if the development of DO was a communal change in these categories, but a generational change in seventeenth century negative declaratives. The second is that the

estimated rate of decline within the individual for each year of his age is well in excess of the rate of increase within the community for each calendar year. Individuals are losing DO much faster than the communal increase, as you can see from the figures in (7): pure generational change should give us a coefficient for age of –1.7. So an account involving generational change would in any case be insufficient.

Given then that the decline with age found in the seventeenth century is indeed age-grading, in the sense of a pattern which repeats stably across generations, the situation is fully consistent with the development of a evaluative set up which differentiates DO NOT from the use of the simple finite with NOT. In the seventeenth century we have less DO in lexically more complex texts, and steep age-grading. This dramatically alters the situation found before 1575, where we have more DO in lexically more complex texts, and no sign of age-grading. But given this, the facts about Ellegård's data (that is, the seventeenth century situation itself and the contrast with the preceding period) are consistent with, indeed strongly imply a differential evaluation which takes effect from the late sixteenth century, whether this is a a hostile evaluation of DO NOT or a positive evaluation of the simple finite with NOT. And it is the onset of this that we see in the sharp drop of DO NOT in Ellegård's database.

References

Biber, Douglas. 1988. *Variation across speech and writing*. Cambridge: CUP.

Cleveland, W. S. 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 74:829-836.

Denison, David. 1993. *English historical syntax: Verbal constructions*. London and New York: Longman.

Ellegård, Alvar. 1953. *The auxiliary* do*: The establishment and regulation of its use in English*. Stockholm: Almqvist & Wiksell.

Kroch, Anthony. 1989. Reflexes of grammar in patterns of language change. *Journal of Language Variation and Change* 1:199-244.

Labov, William. 1994. *Principles of linguistic change: Internal factors*. Oxford: Blackwell.

Nurmi, Arja. 1999. *A social history of periphrastic DO*. Mémoires de la Société Néophilologique de Helsinki LVI. Helsinki: Société Néophilologique.

Rissanen, Matti. 1999. Syntax. In *The Cambridge history of the English language vol III 1476-1776*, ed. Roger Lass, 187-331. Cambridge: Cambridge University Press.

Velleman, Paul, et al. 1995. *DataDesk* (Software for Macintosh). Ithaca: Data Description.

Department of Language and Linguistic Science
University of York
Heslington
York YO10 5DD
England
aw2@york.ac.uk