Published in *Language Variation and Change*, 1989, 1:199-244.

Reflexes of Grammar in Patterns of Language Change[†]

Anthony Kroch

University of Pennsylvania

1. Introduction.

A central methodological problem of historical studies, in linguistics as in other disciplines, is that data are limited to what happens to have survived the vicissitudes of time. In particular, we cannot perform experiments to broaden the range of facts available for analysis, to compensate for sampling biases in the preservation of data or to test the validity of hypotheses. In historical syntax, the domain of this study, the problem is particularly acute, since grammatical analysis depends on negative evidence, the knowledge that certain sentence types are unacceptable. When we study living languages, we obtain such information experimentally, usually by elicitation of judgments of acceptability from informants. Though the methodological difficulties inherent in the experimental method of contemporary syntactic investigation may be substantial (Labov, 1975b), the information it provides forms the necessary basis of grammatical analysis. Hence, syntacticians who wish to interrogate historical material find themselves in difficulty. The difficulty will be mitigated if two reasonable assumptions are made (see, for example, Adams, 1987b; Santorini, 1989): 1) The past is like the present and general principles derived from the study of living languages in the present will hold of archaic ones as well. This assumption allows the historical syntactician to, in the words of Labov, "use the present to explain the past (Labov, 1975a)." 2) For reasonably simple sentences, if a certain type does not occur in a substantial corpus, then it is not grammatically possible in the language of that corpus. Here the assumption is, of course, problematic since non-occurrence in a corpus may always be due to non-grammatical, contextual factors or even to chance. Still, for structurally simple cases, including those we will be considering in this paper, it is unlikely to lead us far astray.

The above perspective on the study of historical materials may be sound, as far as it goes; but, as it is a synchronic perspective, it provides little motivation for pursuing such study, which appears as merely an impoverished version of the synchronic study of living languages. In fact, since data, including negative evidence, are easily obtainable for living languages and since the number and diversity of living languages is so great, one might conclude that historical studies were of marginal significance for general linguistics. They would inform us about the pasts of cultures and give information on the genetic relationships and contacts among languages; but they would not be needed in the construction of a general theory of language.[1] Certain considerations, however, count against such a conclusion. In historical materials, we find a kind of information which is necessarily absent in synchronic data and which offers the prospect of an important contribution to general linguistics from history; that is, information about the time course of language change. With such process information, we may hope to learn how the grammars of languages change from one state to another over time; and from an understanding of the process by which they change, to

---

[†] This paper and the research on which it is based have benefitted over time from the comments and suggestions of a very large number of people. Without hoping to be able to mention them all, I would like to thank particularly the following colleagues and students: Robert Foye, Caroline Heycock, William Labov, John Myhill, Susan Pintzuk, David Sankoff, Gillian Sankoff, Beatrice Santorini, Ann Taylor, and Raffaella Zanuttini. I would also like to thank the many audiences who have heard presentations of the material in the paper. Their insightful comments have very substantially improved it.

[1] This conclusion holds whether we take linguistic theory in a broad or a narrow sense. For example, the study of pragmatics or of discourse structure benefits just as much as syntax from access to experimental evidence from native speaker judgments of acceptability.

learn more about their principles of organization. After all, perturbing a complex system and observing its subsequent evolution is often an excellent way of inferring internal structure. In addition, since the features of any language at a given point in time are the result of a complex interweaving of general principles of language and particular historical developments, knowledge of the historical process by which a language has reached a given state may be important to the proper assignment of responsibility to historical and general factors, and so to the proper formulation of linguistic theory. This paper is meant as a contribution to the enterprise of bringing information on historical process to bear on theoretical issues.

In the discussion which follows, we will present evidence from various linguistic changes that the time course of syntactic change is tightly constrained by the grammar of the changing language. Specifically, we will give evidence that when one grammatical option replaces another with which it is in competition across a set of linguistic contexts, the rate of replacement, properly measured, is the same in all of them. The contexts will generally differ from one another at each period in the degree to which they favor the spreading form, but they will not differ in the rate at which the form spreads. This result, first reported in Kroch (1982 [see Kroch, 1989 for a published version]), is surprising since one might have expected the change to proceed faster in contexts where the advancing form is more common. Indeed, as we will see below, Bailey (1973), in developing his theory of language change, assumes that this must be so, as have other scholars. Our investigations have, nevertheless, developed quantitative evidence for the constant rate hypothesis. In addition, our results show that the grammatical analysis which defines the contexts of a change is quite abstract. We will see that the set of contexts that change together is not defined by the sharing of a surface property, like the appearance of a particular word or morpheme, but rather by a shared syntactic structure, whose existence can only be the product of an abstract grammatical analysis on the part of speakers. Indeed, in some of the cases we will discuss, the competition reflected in the changes under study occurs between entire grammatical subsystems. These competing subsystems have been proposed by syntacticians, on the basis of synchronic analyses, to characterize earlier and later stages of the languages in question, so that the results of our investigation of process will turn out to be consistent with independently motivated structural analyses. In the case we will discuss most fully, the rise of the periphrastic auxiliary *do* in late Middle and early Modern English, the richness of the available database will allow us to see in detail the shaping of the process of change by the grammatical systems in competition.

1.1. Grammar and use.

Studying the process of language change requires a very different method from grammatical analysis. The goal of grammatical analysis, applied to a particular language, is to construct an explicit, finite representation of the set of sentences or sentence types possible in it, where language is conceived, under the familiar idealization, as spoken and learned in an ideal, homogeneous speech community. The goal of grammatical theory is to specify for the set of possible human languages the constraints that all grammars must obey. At the level of both universal and particular grammar, the representations constructed are algebraic and static and, in this respect, have the character of the well-formedness constraints of logical languages. The grammatical perspective provides no vocabulary for the discussion of process. When a language changes, it simply acquires a different grammar. The change from one grammar to another is necessarily instantaneous and its causes are necessarily external. As Saussure put the matter at the outset of modern linguistics:

> [In diachrony] elements are altered without regard for the solidarity that binds them to the whole.... Neither is the whole [system] replaced, nor does one system engender another; one element in the first [i.e., earlier] system is changed, and this change is enough to give rise to another system (Saussure, 1966: 85).

> Language is a system whose parts can and must all be considered in their synchronic solidarity. Since changes never affect the system as a whole but rather one or another of its elements, they can be studied only outside the system (1966: 87).

In the view of contemporary generative linguistics, change occurs between the generations, when children learning a language construct a grammar different from that of their parents (Lightfoot, 1988, 1989) on the basis of altered primary data they are exposed to in the course of acquisition. But how the language data themselves come to change is beyond study, and the fact that documented linguistic changes are generally gradual, with forms slowly replacing one another over centuries, is unaddressable. These processes, therefore, are generally attributed to unanalyzed external factors, usually sociological, like dialect mixture and the conservatism of the written language of most preserved documents.

To study the process of change, we must recognize that the historical texts from which we abstract our data are records of language in use. They have preserved, for us to re-experience or to study, past human linguistic activity; and this activity was not that of ideal speaker/hearers in a homogeneous setting but that of actual people in specific historical circumstances. As sociolinguists have insisted (Weinreich, Labov, and Herzog, 1968; Labov, 1982), people live in linguistically heterogeneous environments and both learn and use their language under these conditions. Furthermore, the widespread occurrence of bilingualism and diglossia show that people often know more than one grammatical system; and the striking phenomenon of intra-sentential code-switching reveals that, in using their knowledge, people may switch fluently between forms from different systems (Poplack and Sankoff, 1981; Woolford, 1983; Joshi, 1985). It is obvious that if we conceive grammar in the standard way, as we have above, the process of language change is not a fact of grammar but a fact of language use and so must be studied with tools appropriate to that domain. The study of language use is the study of the choices that people make among alternative forms in their repertoire of grammatical knowledge in formulating utterances. The usability of grammatical options is sometimes strictly determined by features of extra-sentential context and, to that extent, variation in use may reflect underlying competence extended to the discourse level (Prince, 1988). More germane to our concerns here, however, is another fact, that variation often reflects choices that are not categorically determined by linguistic principles at any level but instead are only probabilistically influenced by features of context and situation.[2] In the case of replacement of one form by another, this is the expected circumstance, at least so long as the change is moving forward and does not turn into a stable alternation. To study such replacement is to determine the nature and weight of these probabilistic factors and to trace their temporal evolution, necessarily using statistical methods applied to corpora of data.

1.2. The time course of language change.

In his book *Variation and Linguistic Theory*, C.-J. Bailey proposes a "wave" model of linguistic change based on two principles (Bailey, 1973). The first is that linguistic replacements follow an 'S'-shaped curve in time, with new forms replacing established ones only slowly in the beginning of a change, then accelerating their replacement in the middle stages of a change and finally, as the old forms become rare, slowing their advance once again. In Bailey's own words,

> A given change begins quite gradually; after reaching a certain point (say, twenty per cent), it picks up momentum and proceeds at a much faster rate; and finally tails off slowly before reaching completion. The result is an ∫-curve: the statistical differences [i.e., differences in frequency of new form versus old form - A.K.] among isolects in the middle relative times of the change will be greater than the statistical differences among the early and late isolects (Bailey, 1973:77).

---

[2] See Bock and Kroch (1989) for further discussion of the stochastic element in language generation.

Bailey's second principle is that differences in the rate of use of a new form in different contexts reflect both the relative time at which the new form began to appear in those contexts and a differential rate of acceptance of that form in those contexts. He says,

> What is quantitatively less is slower and later; what is more is earlier and faster. (If environment a is heavier-weighted than b [i.e., a favors a new form more than b - A.K.], and if b is heavier than c, then: a>b>c.) (Bailey, 1973: 82)

Together these principles constitute a model of the process of linguistic change; and as such they can serve as a jumping off point for the study of specific cases. However, although Bailey enunciates these principles and shows that they can make sense of language variation and change, he does not defend them empirically. He seems, in fact, to think that the principles are too obvious to be contestable, once a proper quantitative examination of change in progress is carried out. Indeed, the first principle does seem to reflect a characteristic property of changes that have been studied quantitatively and the 'S'-shaped curve it refers to will reappear in the cases we present. The second principle, on the other hand, is contradicted by our results, as will become obvious in what follows.[3]

The idea that linguistic changes follow an 'S'-shaped curve is a plausible one, which has been proposed more than once. Thus, we find the suggestion in Osgood and Sebeok (1954) and in Weinreich, Labov and Herzog (1968), as well as in Bailey's book. More recently, Altmann *et al.* (1983) and Kroch (1982, 1989) have proposed a specific mathematical function, the logistic, as underlying the 'S'-shaped curve of linguistic change. While it is not at present possible to demonstrate the correctness of this choice of a specific functional form, its use is generally considered appropriate in statistical studies of changing percentages of alternating forms over time (Tukey, 1977; Aldrich and Nelson, 1984). In the domain of population biology, it is demonstrable that the logistic governs the replacement of organisms and of genetic alleles that differ in Darwinian fitness (Spiess, 1989). Other functional forms which exhibit the 'S'-shape have also been used in studies of changing percentages, most notably the cumulative function of the normal distribution (Aldrich and Nelson, 1984; Piotrovskaja *et al.*, 1974); but these generally differ so little from the logistic that they can provide no improvement in fit to empirical data. Thus, given the mathematical simplicity and widespread use of the logistic, its use in the study of language change seems justified, even though, unlike in the population genetic case, no mechanism of change has yet been proposed from which the logistic form can be deduced. In any case, for the discussion which follows, the choice of an alternative mathematical model would not change the results presented.

The value of proposing a mathematical model for the 'S'-shaped curve of language change is that, using statistical techniques, we can fit data to the mathematical function and estimate its parameters. These estimates can then be compared for different data sets. In this way, data can be used to answer precisely questions concerning the time course of change. The equation of the logistic curve is as given in (1):

$$(1) \quad p = \frac{e^{k+st}}{1+e^{k+st}}$$

where p is the fraction of the advancing form, t is the time variable, and s and k are constants.

An equivalent form is that in (2) where the left-hand side of the equation is the so-called "logistic transform" of frequency (logit):

---

[3] Our results are limited to cases of morphosyntactic change. The extent to which they might extend to phonological changes is unknown.

$$(2)\quad \ln \frac{p}{1-p} = k+st$$

In the equation in (2) we see that the logistic transform of frequency is a linear function of time, so that the two constants of the equation are easily interpretable. The constant s is the slope of the function and hence represents the rate of replacement of the new form by the old, while k, the intercept parameter, measures the frequency of the new form at the fixed point in time, t = 0. For a given value of s, the curve has a fixed form. Changing the value of k merely slides the curve along the time axis. Conversely, changing the time point to which the value t = 0 is assigned alters the value of k, a fact that will become important later in our discussion. Note that because the logistic transform in (2) varies between $-\infty$ and $+\infty$ as p varies between 0 and 1, the logistic, like other functions used in statistics, idealizes the empirical situation. Under the model, there is no time t for which p=0, nor any for which p=1, although as t approaches $-\infty$, p approaches 0 from above and as t approaches $+\infty$, p approaches 1 from below. Of course, actual linguistic changes have starting and ending points, so the model can only approximate real data; and this approximation falsifies the change process precisely at the beginnings and ends of changes. In particular, at the beginning of a change p jumps from zero to some small positive value in a temporal discontinuity which Weinreich, Labov and Herzog (1968) dub the "actuation" of the change.

In principle, actuation might occur in three different ways across the various contexts in which a new form appears. First and most obviously, it might occur sequentially, with the new form appearing at the start in the most favoring context and then successively in less and less favorable contexts. This view is assumed in Bailey's second principle. If the rate of change is at least as high in contexts where the change is actuated earlier than in those where it is actuated later, the change will be more advanced (p will be higher) in the earlier contexts throughout its time course. Presumably, the change will go to completion sequentially as well, as each context approaches and then jumps to a p value of 1. On the other hand, actuation might instead occur simultaneously in all contexts. Then two further possibilities present themselves: either the initial frequency of the new form will be the same in all contexts or it will vary by context. Under the former scenario, at the point of actuation there will be no distinction among more and less favoring contexts, though the rate of increase in the frequency of the new form might differ by context and so create such differences in the course of the change. In the latter scenario, however, the contexts with higher initial p values will begin as favoring contexts and will remain so throughout the change just as long as their rate of change is as high or higher than that of contexts with lower initial p values. As we will see in our discussion of periphrastic *do*, it is possible to distinguish statistically the simultaneous equal activation scenario from the scenarios of sequential and simultaneous unequal activation and to show that this model is not likely to be correct. However, fitting data to the logistic, or to any other function with an infinite domain (hence all those ordinarily used in statistics), will not allow us to choose between the latter two options. One of the interesting features of the data on periphrastic *do* is that other considerations will support the simultaneous unequal actuation scenario against sequential actuation.

Once actuated, a change might spread in a number of ways. It might spread at different rates in different contexts; and if so, it might follow Bailey's second principle. Or, it might spread faster in an initially less favoring context, eventually leading it to become more favoring of the new form than other contexts. In a series of synchronic statistical analyses at different points in time the spread of the change at different rates in different contexts would be reflected in a continual reweighting of the contextual effects on the use of the form.[4] Another option, however, which seems to have been

[4] This reweighting will necessarily occur and will be related mathematically in a simple way to the differences in rate of change among contexts if the synchronic regressions also use the logistic function as their underlying model, as would be the case when the VARBRUL program (Rousseau and Sankoff, 1978) is used. See below for further discussion.

little considered, is that, *contra* Bailey, a change might spread at the same rate in all contexts. Under this option, the differences in the frequency of new forms across contexts could be due either to differences in time of actuation or to an initial difference across contexts established when the change starts which remains constant through time. Because fitting empirical data to the logistic function will allow us to estimate the slope parameters for each context of a changing form, we can determine, where sufficient data are available, whether the rates of change in different contexts are the same or different, putting the part of Bailey's principle that makes a claim about rates of change to an empirical test. Here, as we have already mentioned, the results to be presented give clear support for the hypothesis that changes spread at the same rate in all contexts and as we will see, on the basis of this constant rate hypothesis substantial progress can be made in understanding the relationship between the structural patterns uncovered by grammatical analysis and the frequency patterns revealed by sociolinguistic methods.

2. Evidence for the constant rate hypothesis.

In this section, we present evidence from specific historical changes that have been studied quantitatively to support our constant rate hypothesis. As stated above, although the rate of use of grammatical options in competition will generally differ across contexts at each period in time, the rate of change will be the same across contexts. In some of the cases that we will examine, the original studies trace the developments context by context, and the rate of change for each can be estimated directly. In others, however, multivariate analysis with the VARBRUL program (Rousseau and Sankoff 1978) is used; and here our result appears in a different form. A constant rate of change across contexts is mathematically equivalent to fixity of contextual effects, in direction and size, across time periods. Thus, if a study reports a series of multivariate analyses for different time periods, and the contextual effects are constant across these analyses, the rate of change of each context measured separately would necessarily be the same. This equivalence holds because, in statistical terms, the constant rate hypothesis is the claim is that the overall rate of use of a form is independent of the contextual effects on its use. This statistical independence can be expressed by modeling usage of the old and new forms undergoing replacement in a set of equations of the following form, one for each of the combinations of contextual features that occur in the language:

$$(3)\quad \ln \frac{p_C}{1-p_C} = f(t) + a_1 + a_2 + \ldots$$

where p_C is the probability of the new form being used in context C, f(t) is a linear function of time, and the a_i's are constant weights associated with each contextual feature of context C, positive for favoring features and negative for disfavoring ones.

The equations of (3) are, of course, those used by the VARBRUL program for estimating factor effects for variable rules, and f(t) can be taken to represent the "input probability" of a VARBRUL analysis as a function of time. As is clear from the equation, the contextual effects are constant across time and do not interact with the time variable.

Our evidence for the constant rate hypothesis comes from four different quantitative studies, three by other researchers and one of our own. The studies by others are: 1) the account by Noble (1985) of the increasing use of *have got* in place of *have* to express possession in British English between the eighteenth and twentieth centuries; 2) the description by Oliveira e Silva (1982) of the increasing tendency in continental European Portuguese to use the definite article in possessive noun phrases; and 3) the quantitative analysis by Fontaine (1985) of the loss of subject-verb inversion (the verb-second constraint) in the history of French. Our own study is a re-analysis of the history of periphrastic *do* based on the well-known and extensive description in Ellegård (1953) and our own previous work (Kroch, 1982; 1989). It is presented in detail in section 3 below. Of the other studies, the first two describe morphosyntactic replacements that apparently do not

implicate broad subsystems of the grammar and will, therefore, be discussed only briefly. The loss of the verb-second constraint in French, on the other hand, involves a large-scale reorganization of the syntax of French and requires a more extensive treatment.

2.1. The replacement of *have* by *have got* in British English.

In the course of the past three centuries there has been a regular drift toward the replacement of main verb *have* by the idiomatic form *have got* in British, and to a lesser extent in American English. The alternation is illustrated in the following pair of sentences cited in Noble (1985):

(4) a. Anyhow, she has what amounts to a high Cambridge degree. (1898)
b. You've got plenty of hair. (1968)

Choice of the *have got* form avoids use of *have* as a main verb, which in British English until recently preserved an auxiliary-like syntax. As such, the rise of *have got* may a late continuation of the linguistic changes surrounding the rise of *do* to be discussed later. At the moment, however, we have little evidence to support such a link and will not pursue the matter further.[5] Instead, we will focus exclusively on the results of Noble's quantitative investigation as it bears on our main hypothesis. Noble collected all examples of main verb *have* and *have got* where an alternation between the two is grammatically possible from a corpus of British and American plays and other materials likely to show linguistic usage of the spoken language, spanning the period from 1750 to 1935. Only the British data are analyzed in sufficient detail to be useful for our purposes. In her British corpus, she traced two major contextual effects on the competition across time based on the semantic character of the possession. She found that whether the possession was temporally bounded or permanent and whether the possessed element was a concrete object or an abstract quality had a large effect of the choice of verb form. When the possession was temporally bounded or the possessed element was a concrete object, the use of *have got* was favored, not a surprising result given that *got* is historically, and in British English to this day, the past participle of *get*, whose meaning is "to acquire". These semantic considerations, however, only affected the rate of use of *have* versus *have got* and did not determine categorically which form was chosen. Thus, in the following examples, both forms are possible in all cases:

(5) a. I've got / I have a new job. [temporally bounded]
b. I've got / I have brown eyes. [permanent]

(6) a. She's got / she has a car. [concrete object]
b. She's got / she has a careful approach. [abstract quality]

Tables 1 and 2 present Noble's results:

Period	type	% *have got*	total	prob.
1750-1849	tempor. b.	12	83	.66
	permanent	4	108	.34

[5] Noble does report some evidence to support a link between use of *have got* and avoidance of main verb *have* for its auxiliary-like syntax. She finds that in British English the use of *have got* is favored in negative sentences and questions, where the auxiliary-like behavior of main verb *have* is manifest. This effect is found only in British English and not in American English. This is the expected result since the use of *do* support with *have* is found much earlier and more extensively in the American dialect. Once *do* support extends to main verb *have* it no longer has auxiliary-like syntax.

1850-1899	tempor. b.	34	99	.64
	permanent	16	122	.36
1900-1935	tempor. b.	89	74	.66
	permanent	70	43	.34

Table 1: Effect of possession type on the choice between *have* and *have got*.

Period	type	% *have got*	total	prob.
1750-1849	concrete	13	68	.66
	abstract	4	123	.34
1850-1899	concrete	34	74	.61
	abstract	20	147	.39
1900-1935	concrete	86	51	.58
	abstract	79	66	.42

Table 2: Effect of concreteness on the choice between *have* and *have got*.

The results here are as expected. Looking at the probabilistic weights in the righthand columns of the tables, we see that the effect of temporally bounded versus permanent possession favors the former by approximately .65 to .35 in all three periods and that of concrete object versus abstract quality favors concreteness by approximately .60 to .40 across the time periods.[6] Noble does not report the statistical significance of the small differences from period to period, but these results are as close to predicted as one could hope for, given the small amount of data analyzed and the broad width of the time periods.

2.2. The rise of the definite article in Portuguese possessive noun phrases.

In Modern Portuguese, noun phrases with prenominal possessive pronouns[7] are usually introduced by the definite article, as opposed to in English, where noun phrases like "*the my book" are impossible. The examples below illustrate the variability of article use:

(7) a. Maria conhece (o) meu irmão.
M knows the my brother
b. (O) meu unico cão morreu.
the my only dog died
c. (Os) seus livros são interesantes.
the his books are interesting

In early Portuguese, however, the use of the definite article in such noun phrases was impossible. Figure 1 below, based on data in Oliveira e Silva (1982), shows the rise in the frequency of the form in Continental Portuguese, the dialect where the replacement is most regular:

---

[6] In the version of the VARBRUL program used by Noble the average probabilistic weight value of factors in a group of contextual alternatives is constrained to equal .5. Values above .5 favor success (here use of *have got*) and values below .5 disfavor it, by comparison to the overall rate of success for the whole sample.

[7] Unlike English, Portuguese does not admit prenominal possessive full noun phrases, as in "John's car."

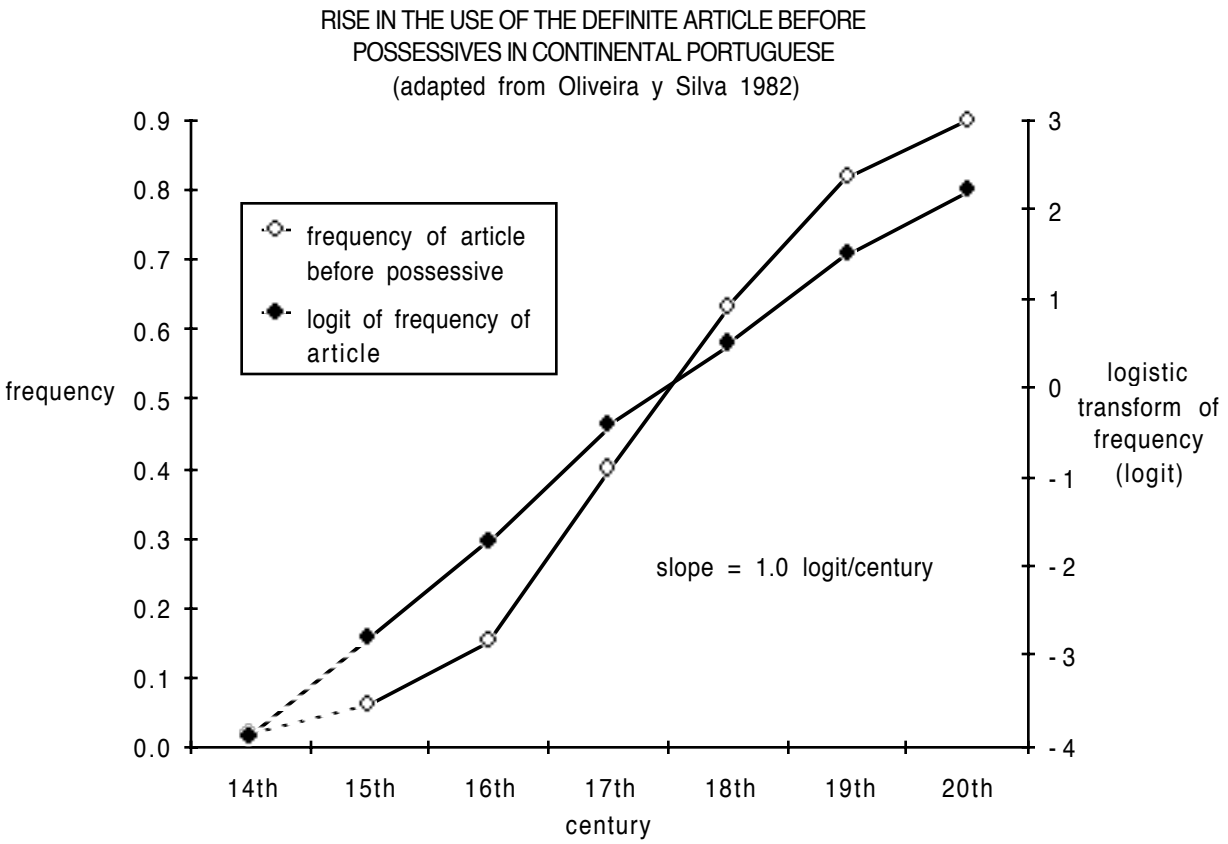


Figure 1: Rise in the use of the definite article before possessives in Continental Portuguese (adapted from Oliveira e Silva 1982)

The curve with the open diamonds represents the percentage of article use and the one with the black diamonds the logistic transform of the percentages (see equation (2) above). The former exhibits the characteristic 'S'-shape of linguistic change, and latter is very close to a straight line, showing that the logistic function fits these data well. Oliveira e Silva found 4 factors of the pragmatic and morphosyntactic context of noun phrases that significantly influenced the use of the definite article with a following possessive pronoun: whether the possessed noun is a kinship term, whether the possessive noun phrase has a unique referent in the discourse context, whether the possessive noun phrase is the object of a preposition, and whether the possessive pronoun was third person.[8] Using the VARBRUL program, she estimated each of these effects for data from the six centuries from the fifteenth to the twentieth and obtained the results in Figure 2:

[8]Discussion of other effects investigated by Oliveira e Silva, particularly the morpho-phonological, is beyond the scope of this article.

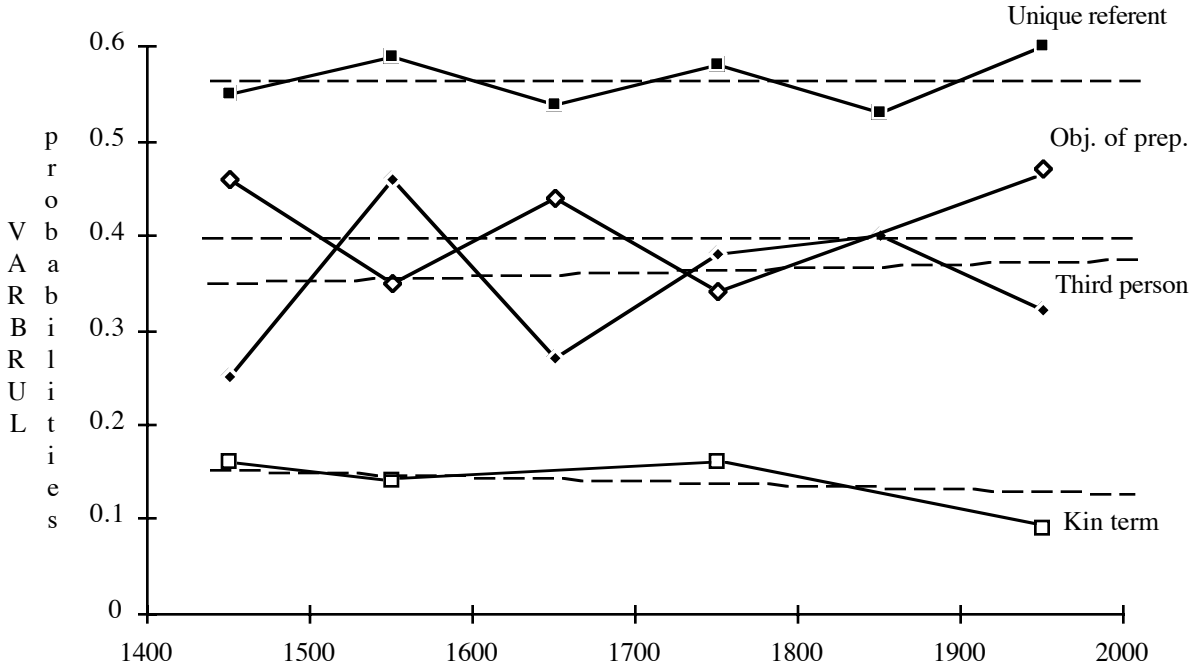


Figure 2: Stability of factor effects over time on the use of the definite article with possessives. (adapted from Oliveira e Silva 1982)

As Oliveira e Silva notes, the factor effects are remarkably stable here. The dashed regression lines for the various factors against all show slopes very close to zero; that is, there is no change with time in the effects of these factors. In other words, her results show that the increasing use of the definite article in Continental Portuguese increases at the same rate in all contexts. As in the case of Noble's work, Oliveira e Silva gives us no measure of the statistical significance of the contextual constancy effects she has found; but the results are again quite strong, given the small size of her samples for each time period.

2.3. The loss of verb-second word order in French.

Fontaine (1985) presents a quantitative description of the loss of subject verb inversion in Middle French whose results are directly relevant to our hypothesis; and recent work by Adams (1987a, 1987b) on the grammatical analysis of Old French and on the character of the syntactic change in Middle French makes Fontaine's quantitative results even more interesting for us.[9] Adams points out that Old French was a verb-second language; that is, when a noun phrase complement was topicalized, the tensed verb of the sentence had to appear as the immediately following constituent. This requirement is identical to that observed by the modern Germanic languages, with the exception of English. In (8) below, we give examples of verb-second word order:

(8) a. Messe e matines ad li reis escultet. (Adams, 1987b:2)
mass and matins has the king heard

[9] In our discussion we will rely on the analysis in Adams (1987a, 1987b), though we recognize that many aspects of her analysis remain controversial. See, in particular, the articles in (Hirschbühler and Rochette, 1988) for detailed discussions of the grammatical issues.

b. Moi doiz tu dire ton afere. (Adams, 1987b:147)
to me must you tell your affair

The verb-second constraint appears to be variable in sentences with preposed adverbs and prepositional phrases. In some cases, the adverb counts as the trigger for fronting of the verb, but in others it does not. The two word order possibilities are illustrated in (9):

(9) a. Or a Nostre Seigneur tout-à-coup fait cheoir si grand et somptueux édifice.
Now has Our Lord suddenly made arise so grand and sumptuous edifice.
(Adams, 1987b:90)
b. Dejuste lui li dux Neimes chevalchet. (Adams, 1987b:192)
Beside him the duke N rode

This variability, however, is generally agreed not to lie in the applicability of the verb-second constraint but rather in the placement of the fronted adverb. Sometimes it is preposed to the position of topicalized complement noun phrases, where it triggers verb-fronting, and sometimes to a position before the topicalization position, yielding a superficially verb-third sentence. In this latter case, the subject usually serves as the verb-second trigger. This variable positioning of preposed adverbs and prepositional phrases is quite general in the medieval Germanic languages, which are otherwise always verb-second, though it has largely disappeared from the modern languages.

Another complication in the verb-second pattern of Old and Middle French reflects the special character of French pronouns. From Old French on, unstressed object pronouns were preverbal clitics, and they did not count as occupying a syntactic position for the verb-second constraint. Thus, the example in (10) would not count as an exception, and indeed could have no other word order with an unstressed object pronoun:

(10) Tu m'as amé celéement et jou toi. (Adams, 1987b:103)
you me-have loved secretly and I you

In Middle French, subject pronouns became clitics and were variably ignored for the purposes of the constraint. Thus, the word orders in both (11) and (12) were possible in Middle French:

(11) a. ...en riant tant que à peine le pouvait-il dire.
in laughing so much that barely it-could-he say
(Hirschbühler et Junker, 1988:79)
b. ...si ay je moi. (Adams, 1987b:183)
so have I myself

(12) a. Et ce conseil nous vous donnons. (Adams, 1987b:195)
and this advice we you give
b. Car grant voulenté ils avoient de faire fais d'armes. (Adams, 1987b:195)
because great desire they had to do deeds of-arms

A final characteristic of the verb-second pattern in French is the distribution of sentences with null subjects. Unlike the Germanic languages, both Old and Middle French allow null subjects in every grammatical person. Adams argues, however, that these null subjects occurred only in inverted position; that is, in sentences with some non-subject constituent preposed; and when the verb-second constraint disappeared from the language, so did null subjects (but see Vance, 1988). Examples of null subject sentences are given in (13):

(13) a. Or voi ge bien, plains es *pro* de mautalant. (Adams, 1987b:6)
now see I well, full are (you) of ill-will

b. De venoison ont *pro* grant plenté. (Adams, 1987b:155)
Of meat have (they) great plenty

Verb second word order, as measured by the rate of inversion of subject with tensed verb after preposed constituents disappeared from French in the course of the Middle French period, between the fourteenth and the sixteenth centuries. Given Adams's analysis, the constant rate hypothesis leads us to expect inversion to decline at the same rate as the use of null subjects. The quantitative data collected and analyzed by Fontaine allow us to test this prediction. Figure 3 (Fontaine's Figure 10) directly shows the decline of inversion and of null subject use plotted as the logistic transform of the frequency of each.

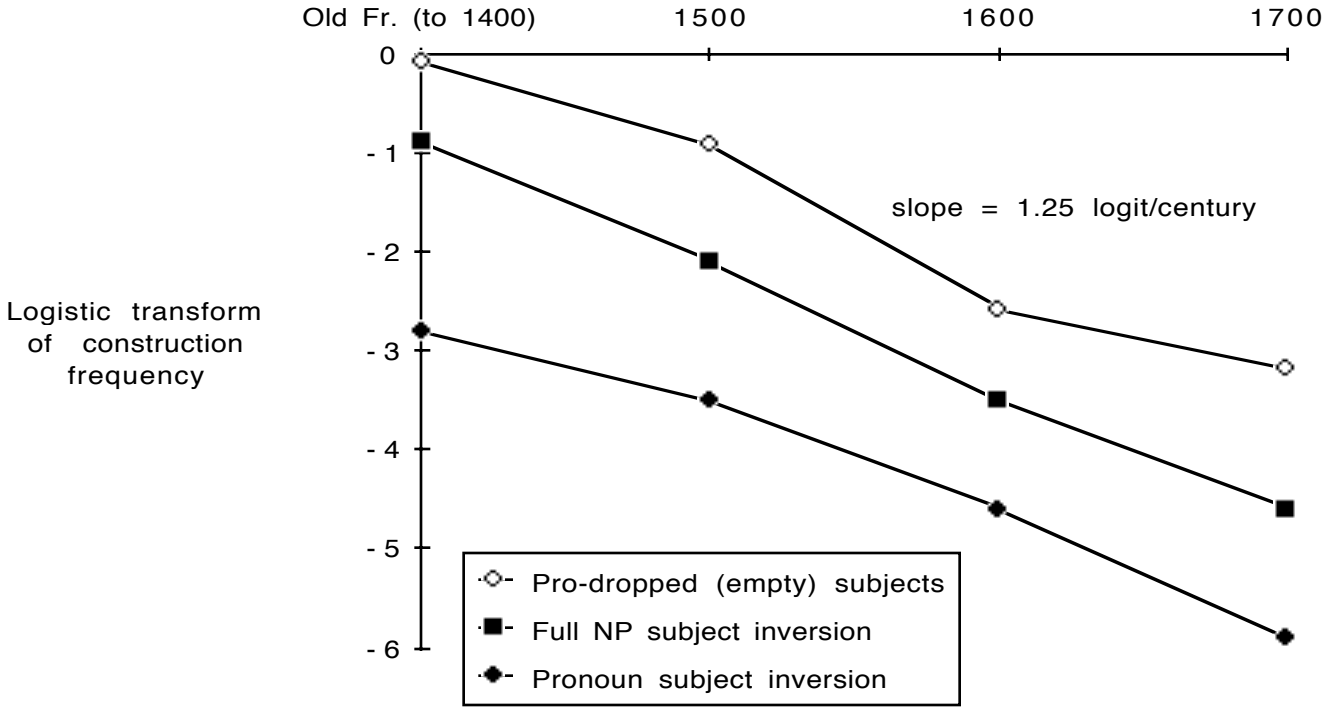


Figure 3: The decline of inversion and null subjects in Middle French.
(adapted from Fontaine 1985:90)

The curves in the graph are nearly straight lines, indicating that the logistic function fits these data well, as in the Portuguese case. Furthermore, the rates of loss are very close to one another. Again, we do not have significance figures for the small observed differences in slope across contexts but it is clear by inspection that the constant rate hypothesis is supported by these data.

The most significant feature of Fontaine's results is that they show the grammatical analysis unifying the loss of inversion and the loss of null subjects to be quite abstract. The change cannot be conceived as a competition among surface forms, say inverted and non-inverted word order. Rather, it is the entire grammatical complex that enforces the verb-second contraint which is in competition with the Modern French SVO (subject-verb-object word order) grammar. This point can be strengthened and more evidence provided for the constant rate hypothesis if we pursue the grammatical analysis of the change and develop further a line of argument in Adams (1987b), though in a way not entirely consistent with her discussion. Adams points out that the loss of the verb-second constraint is associated with a change in the phrasal accent pattern of French and

argues that the change in accent may have played a causal role in the word order change. While Old French, like the Germanic languages, allows accents both phrase initially and finally, Modern French allows only a single, phrase-final accent. The accentual difference is reflected in differences in the syntax of topicalization in Modern French as opposed to Old French, English and German. In the latter languages, topicalized non-subject constituents receive a secondary accent in addition to the primary accent that falls on the clause from which they are extracted. For example, in a sentence like (14), both *beans* and *adore* receive stress:

(14) I can't stand squash, but bèans$_i$ I adóre e$_i$.

By comparison, the French sentence corresponding to (14) is ill-formed. Instead of the topicalization (15a), French requires an apparent left dislocation with a resumptive clitic pronoun in the thematic position of the preposed constituent:

(15) a. *Je déteste les courgettes, mais les haricots$_i$, j'adore e$_i$.
b. Je déteste les courgettes, mais les haricots$_i$, je les$_i$ adore.

One possible reason for the use of left dislocation in French is that topicalization requires two stresses for semantico-pragmatic reasons (Prince, 1984). In French this double stress pattern can only occur if the preposed constituent is in a separate major intonation phrase from its clause; i.e., if it is left-dislocated and binds a pronoun in the clause. Supposing Adams to be correct, one might hypothesize that the loss of verb-second word order in French took place via the replacement of topicalization by left-dislocation. After all, as we have already noted, preposed adverbs and prepositional phrases may hold two positions in Old French, one the topicalization position and another further to the left, which is syntactically indistinguishable from the position of left dislocation. Even in Old French, left dislocation of noun phrases was possible, if rare, as the following example shows:

(16) [Cels qu'ils unt mort]$_i$, ben les$_i$ poet hom priser. (Priestley, 1955:10)
those who they have killed, well them could one praise

Suppose then that the change in phrasal accent forces preposed constituents to move from the topicalization position to the position of left dislocation. The result will be that the preposed constituents no longer function as verb-second triggers and sentences will appear to be verb-third. However, they will not violate the verb-second grammar since left-dislocated elements do not count for the verb-second constraint (witness example (16)). The topicalization position will be filled by the subject of each sentence, the only noun phrase that can ordinarily be topicalized without stress in verb-second languages. In sentences with preposed adverbs and prepositional phrases, the only effect of the change in accent on word order will be a decline in the rate of subject-verb inversion; but in sentences with preposed noun phrase complements, there will be an additional effect, an increase in the rate of use of the resumptive clitic pronouns required by left dislocation. This difference is illustrated in the following pair of Modern French examples:

(17) a. À Jacques$_i$, le professeur parle e$_i$ tous les jours.
to J the teacher speaks every day
b. Les autres arts et sciences$_i$, Alexandre les$_i$ honoroit bien.
the other arts and sciences, A them-honored well
(Priestley, 1955:21)

Over time, the number of sentences which provide positive evidence for the verb-second constraint will decline relative to those, like (17), which are also consistent with a simple SVO grammar. Eventually, the absence of sufficient positive evidence will trigger a grammatical reanalysis and subject-verb inversion will no longer be possible.

Applied to the study of the time course of the change in French word order, the above analysis predicts that the rate of left dislocation should rise as the rate of inversion falls; and the constant rate hypothesis further predicts that these two rates should be identical. Fortunately, these predictions can be tested. A quantitative study by Priestley (1955) looked at the rate of use of left dislocation, which he called pronominal "reprise", through the history of French. Figure 4 summarizes his results:

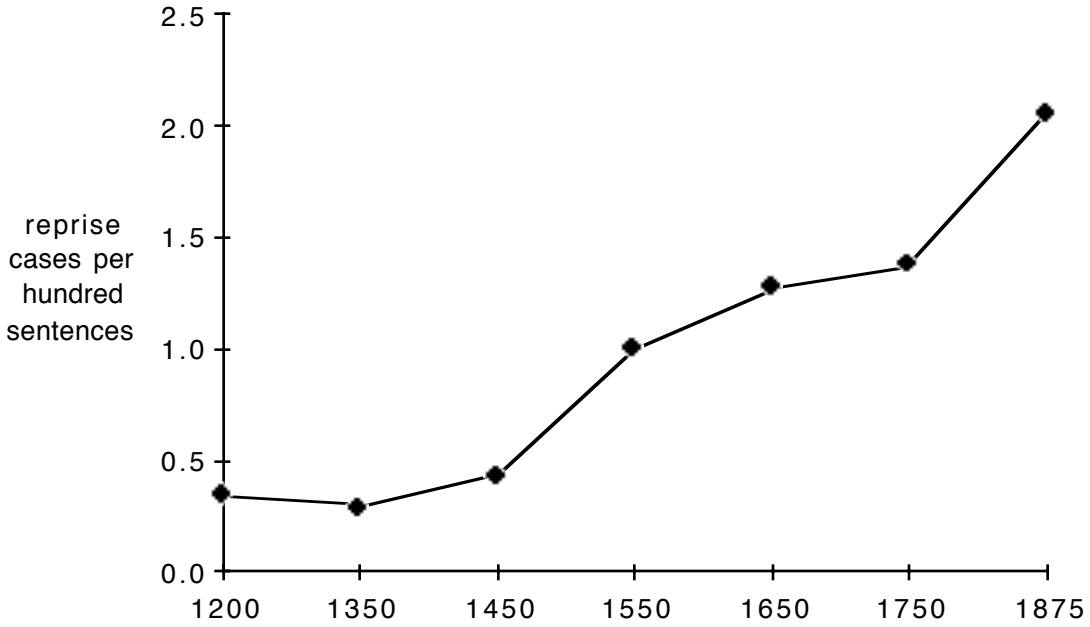


Figure 4: Rise of reprise (left dislocation) with preposed direct objects in French prose (Priestley, 1955).

If we fit a logistic curve to Priestley's data via regression[10] and compare the logistic transform of the fitted curve with Fontaine's results, we obtain the pattern in Figure 5 below. We have, of course, reversed the sign of the slope of the regression, since the rise in left dislocation corresponds to the loss of topicalization. It is important to note that the data for Fontaine's curves are, in the vast majority, sentences with preposed adverbs and prepositional phrases, not noun phrases. This is due to the rarity of noun phrase preposing in texts by comparison to the fronting of adverbs and prepositional phrases. Thus, in comparing Priestley's data to Fontaine's, we are not simply looking at the same type of data twice.

[10] We assume that the change in frequency of reprise begins in 1350 and that the end of the change is reached in 1750. The sharp jump from 1750 to 1875, we take to be caused by another effect. From Priestley's data, it would appear that this rise is caused by an increase in the left dislocation of subjects, which is, of course, irrelevant to our analysis.

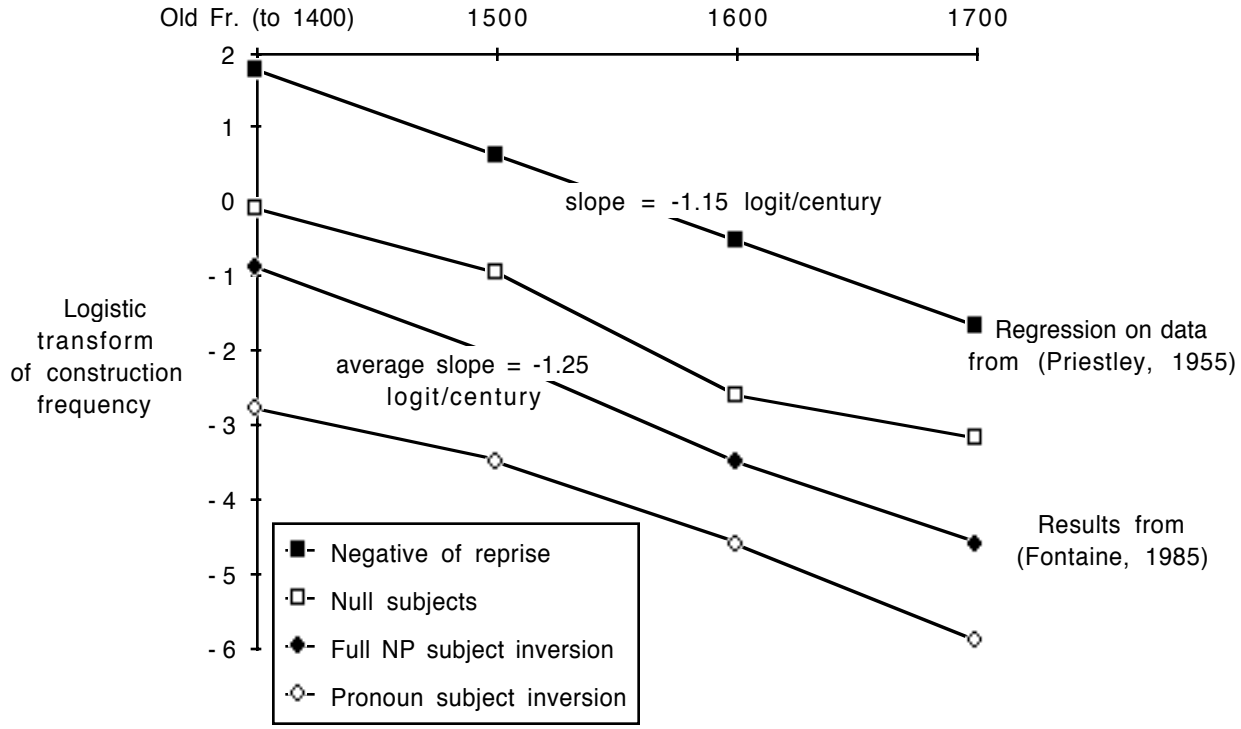


Figure 5: Decline of inversion and null subjects in Middle French compared to the rise in reprise.

Once again, we find the expected pattern. The rise in reprise proceeds in tandem with the loss of inversion and at the same rate. The data thus provide at once quantitative evidence to support our grammatical analysis and further support for the constant rate hypothesis.

3. The case of periphrastic *do*.

One of the most extensively studied syntactic changes in the linguistic literature is the rise of periphrastic *do* in English questions and negative sentences (Denison, 1985; Engblom, 1938; Ellegård, 1953; Kroch *et al.* 1982, Kroch 1989; Stein, 1986). In Middle English, questions were formed by inverting the tensed verb with the subject and negative sentences by placing the negative marker *not* immediately after the tensed verb, as in (18) and (19), respectively:

(18) a. How great and greuous tribulations suffered the Holy Appostyls...? (302:166:10)[11]
b. and in thy name have we not cast oute devyls...? (310:31:45)

(19) a. ...spoile him of his riches by sondrie fraudes, whiche he perceiueth not. (346:86:23)
b. Go, say to hym we wyll not grefe [grieve]. (218:8:292)

[11] Many of the examples in this paper are taken from the sources in (Ellegård 1953). All such sources are identified with Ellegård's numbering system, as follows: (source number: page number: line number). The key to these sources may be found in the appendix to Ellegård's work.

In sentences with auxiliary verbs, like the (b) examples above, the Middle English pattern is preserved in Modern English; but, beginning some time in the 14th century, the pattern began to change in the case of sentences with a tensed main verb, with *do* appearing as a dummy auxiliary, as in (20):

(20) a. Where doth the grene knyght holde hym? (304:97:15)
b. ...bycause the nobylyte ther commynly dothe not exercyse them in the studys therof. (318:194:567)

By 1700, this new form had largely, though not entirely (see Rydén, 1979), replaced the original usage. Interestingly, until the middle of the 16th century, the increasing use of periphrastic *do* in contexts like (20) was accompanied by an increase in its use in ordinary affirmative declarative sentences like those in (21):

(21) a. They worschipped the sonne whanne he dede arise. (78:327:8)
b. When he dyd se[e] that Crist schold be dede.... (167:188:2)
c. Me thinke I doe heare a good manerly Begger at the doore...(346:5:17)

In the modern language, the use of auxiliary *do* in uninverted affirmative declaratives is limited to emphatic contexts like (22), where its emphatic character is demonstrated by the fact that it bears stress:

(22) a. John says he doesn't like pizza but he dóes like it.
b. You haven't been here in a long time so when you dó come, be sure not to lose your way.

It is evident, however, both from the frequency of examples like (21) and from the discourse context of many of them (Ellegård, 1953) that unemphatic affirmative declarative *do* was in common use in the middle of the sixteenth century, after which its use declined until the modern situation was reached. Only in archaic legal phrases like, "I do hereby declare..." does the modern language still exhibit unemphatic *do* in this context.

3.1. The grammatical character of the change.

The rise of periphrastic *do* is contemporaneous with the development of the class of syntactically distinctive modal auxiliaries in English, and according to Lightfoot (1979), these two phenomena are reflexes of the development of the category 'auxiliary verb' in English. In Old and early Middle English the modals could be main verbs, as in Modern German. They could take direct objects and they could appear as the complements to other verbs. They also exhibited an extensive pattern of person and number agreement, although the third person singular ending was zero rather than -s/-th for the class of verbs to which they belonged (the preterit-present verbs). In the course of later Middle English, contemporaneously with the rise of *do*, the modals lost their status as main verbs and became syntactic auxiliaries, obligatorily appearing as the highest and tensed verb of their clause. Lightfoot (1979) argues that this change amounted to the creation of a new syntactic category in English. He sees it as a consequence of the general decline in verbal inflection. Once the plural and infinitive endings were lost, the modals only inflected for person in the second person singular ("thou canst", etc.). This weakness of inflection was further accentuated by the fact that present and past tense forms of the modals were from early times not related to each other in the standard way, a fact still characteristic of current usage, where "John may go" and "John might go" ordinarily do not differ in tense. Given the subjunctive-like semantics that the epistemic modals characteristically exhibit, it was then not surprising that they should become syntactic auxiliaries restricted to the position where tense and mood inflection appear. Roberts (1985) argues that a crucial factor in the development of the modal auxiliary was the collapse of the

subjunctive mood in the course of Middle English, which led to the use of modals in place of the subjunctive inflection (see also Steele *et al.*, 1981). According to Roberts, when used as replacements for the subjunctive, the modals could not be verbs taking noun phrase arguments but rather had to be operators on clauses. Such operators, tense and mood, occur under the INFL(ection) or AUX(iliary) node in a conventional phrase structure tree; and so, as operators, modals would be generated under that node. Under Roberts' view, it is the epistemic use of modals which is ultimately responsible for the change in their syntactic positioning, a point to which we will return. Warner (1982, 1983) gives evidence that some modals were already exclusively auxiliary verbs much earlier than Lightfoot claims, pointing out that the modals *must* and *shall*[12] have no attested occurrences as other than the highest and tensed verbs in Middle English and perhaps not even in Old English. Thus, the category 'auxiliary verb' must have existed throughout Middle English; and the change described by Lightfoot was not the creation of such a category but the lexically conditioned drift of the other modal verbs into the category in the course of time. This account is consistent with the situation in Modern German, where the modals, when they function semantically as mood-bearing auxiliaries, must be tensed, even though, in their root senses, they can function as main verbs and can appear as non-finite complements to other verbs (Kroch, Santorini, and Heycock 1987). In any case, the modals' transition to auxiliary status seems to have been completed by the middle of the sixteenth century. Although certain root modals, particularly *can*, continue to appear sporadically as main verbs and as complements to other auxiliary verbs into the nineteenth century (van Kemenade, 1989), these occurrences are rare and should be treated as occasional archaisms. For our purposes, the crucial fact concerning the modals is that, like the verbs *have* and *be*, they never appear as complements to periphrastic *do*. This fact argues that, from its first appearance, periphrastic *do* subcategorizes only for verbs that never function as auxiliaries and hence that the categorial distinction auxiliary/non-auxiliary must have been available from the beginning of the Middle English period, when the first instances of periphrastic *do* appear. Thus, like Warner and other recent commentators, we reject the analysis in Lightfoot (1979) (largely abandoned in Lightfoot, 1988) that the category auxiliary verb appears in English only at the end of Middle English.

A more plausible scenario relating the rise of periphrastic *do* to the development of the Modern English auxiliary verbs can be developed on the basis of the account in Roberts (1985). As Roberts points out, the crucial syntactic distinction between main verbs and auxiliaries in Modern English is that the former, unlike the latter, cannot appear in the INFL position of a phrase structure tree. In this respect, Modern English differs both from other modern western European languages and from Middle English. In these languages, if a sentence contains no auxiliary verb, the main verb raises from its base position as the head of the V(erb) P(hrase) to the INFL position (Emonds, 1978; Koopman, 1984). In French, to take a well-known example, the placement of weak sentence adverbs, of negation, and of so-called "floated" quantifiers marks the boundary between INFL and the VP. Thus, in sentences with auxiliary verbs, these elements are ordinarily found between INFL and VP. The examples in (23) illustrate the pattern:

(23) a. Les hommes sont toujours partis avant midi.
the men are always left before noon
b. Les hommes (ne) sont pas partis.[13]
the men neg1 are neg2 left
c. Les hommes sont tous partis.
the men are all left

---

[12] The evidence regarding *shall* was incomplete at the time Warner wrote his article, because the Middle English Dictionary had not yet reached the letter 's'. Since then, the Dictionary entry for *shall* has appeared; and as Warner expected, there are no instances of it functioning as a main verb.

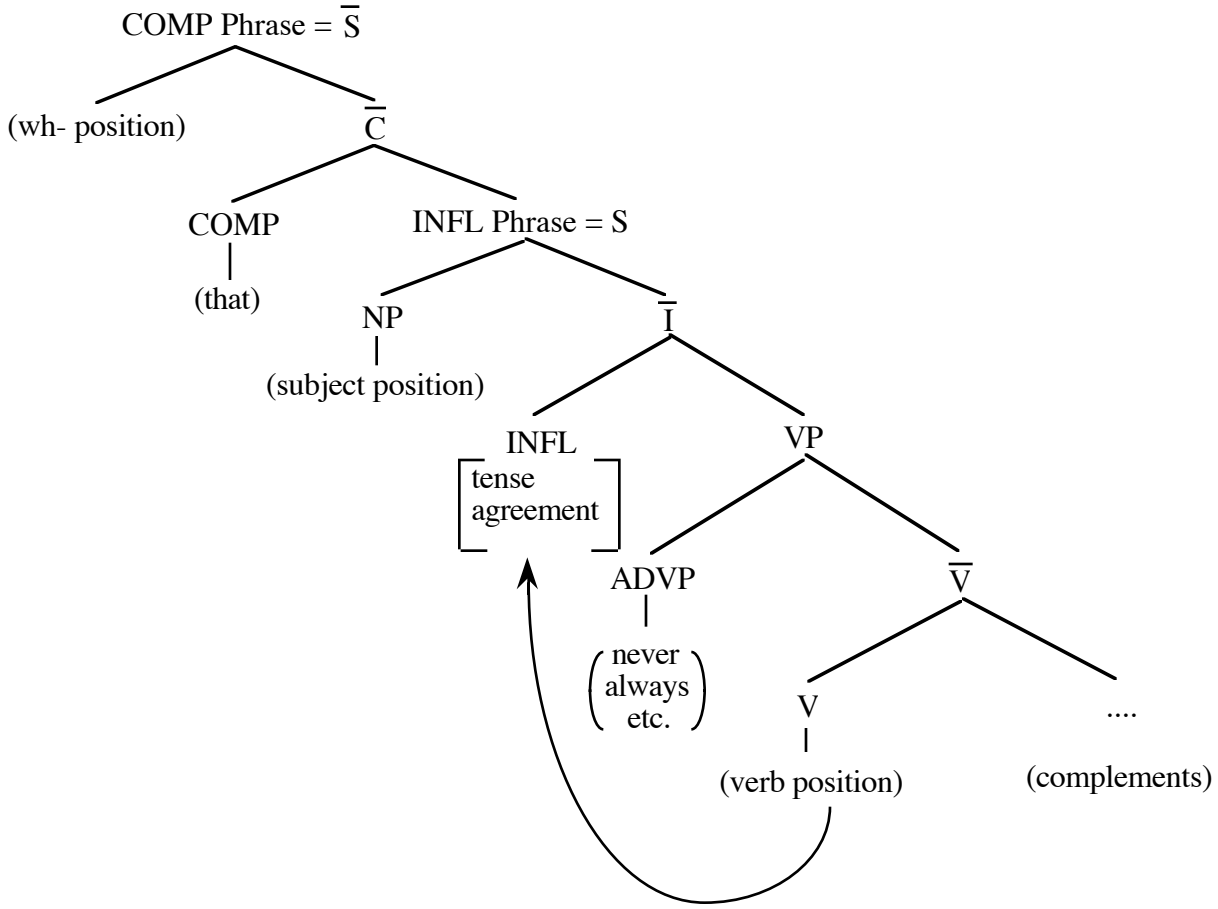
[13] In French there two negative morphemes that together mark sentence negation. The first, which may be deleted, is *ne* and it comes immediately before the tensed verb. The second, *pas*, comes between INFL and VP and is the one of interest to us in this discussion.

When the sentence contains no auxiliary verb, however, these elements are found after the main verb:

(24) a. Les hommes partent toujours avant midi.
b. Les hommes ne partent pas.
c. Les hommes partent tous.

Transformational grammarians account for this pattern by postulating the rule of V(erb)-to-I(NFL) raising, which moves the highest verb in underlying structure to the INFL node. When the sentence contains an auxiliary verb, the highest verb structurally will be the auxiliary; but when there is no auxiliary, it will be the main verb. The raising is obligatory in tensed sentences because only through it can the finite verb be marked with tense, mood, and agreement, which are syntactic features or morphemes generated under INFL. The best evidence for the rule is that it accounts neatly for the permuted word order of main verb and adverb-like elements illustrated in (23) and (24). The underlying structure and movement are shown graphically in the following tree, drawn using the phrase structure conventions of Chomsky (1986):

(25)



Middle English behaved like French with regard to the positioning of adverbs, negation and floated quantifiers[14], as illustrated below:

(26) a. It shal wel abyde and tary for me. (Caxton, *The Ryall Book*, line 20-25 [excerpt in Mossé (1952)])
b. Swylke ('such') kane noghte fyghte for thaire hony... (Rolle, "The Bee and the Stork," lines 24-25 [excerpt in Mossé (1952)])
c. [Januarie] wolde bothe assayen his corage in libertee and eek in mariage; (Chaucer, *Merchant's Tale*, lines 1724-25)

(27) a. I wende wel thys nyght to have deyed ('I managed almost tonight to die') (Caxton, *The Ryall Book*, line 25-30 [excerpt in Mossé (1952)])
b. ...if thay do noghte all as they wolde till ('to') tham. (Rolle, "The Bee and the Stork," lines 23-24 [excerpt in Mossé (1952)])
c. ...that is to seyn whil that they lyven both ('while they both live') (Chaucer, *Parson's Tale*, lines 915-920)

Therefore, we may suppose that its grammar also contained the rule of verb raising to INFL. Modern English, in contrast, must lack this rule since weak sentence adverbs, negation, and floated quantifiers never appear after the main verb. In sentences with auxiliaries, the placement of these elements is the same as in Middle English:

(28) a. John has always liked bananas.
b. John isn't eating today.
c. The men will all leave together.

But in sentences with only a main verb, the order fails to permute as it did in the earlier language:

(29) a. John always eats after 8 o'clock.
b. John doesn't know much about cars.
c. The men all complained about the wages.

Following other treatments, beginning with Chomsky (1957), Roberts proposes that in Modern English the tense marking of main verbs occurs via a transfer of the affix from INFL to the verb in its deep structure position ('affix hopping'). This transfer is blocked by the sentence negator *not*, and periphrastic *do* is inserted to provide a lexical support for the affixes in INFL. The appearance of *do* in questions is also motivated by the need for a lexical support for the affixes. Both Middle and Modern English form questions by preposing the tensed verb. In Middle English this verb could be either an auxiliary or a main verb, but in Modern English it must be an auxiliary. Thus, we have the contrast between (18a) and (20a) above or between (30) and (31) below:

(30) Thynkest thou to avoyde that neuer mortall creature might escape? (302:192:43)

(31) ...dyde he begyle us that sayd they were spoken of the[e]? (302: 219-33)

A widely accepted analysis of these facts is that English questions have always been signaled by the fronting of INFL to the COMP(lementizer) position (see (25) above), the position where introductory 'that' appears in subordinate clauses. In Middle English this fronting occurred subsequent to verb raising to INFL; and in Modern English the appearance of the subject between INFL and the main verb blocks transfer of the affix, forcing *do* insertion. This analysis is supported by the fact that *do* insertion is not found in questions where INFL and the main verb

[14] The case of quantifier floating is more complex than the others, and for reasons not relevant to our concerns there are few examples in Middle English parallel to the French cases (see Lightfoot, 1979).

remain adjacent after INFL is fronted to COMP. In questions signalled by intonation alone, *do* does not appear, nor does it appear in examples like (32), where the subject of the sentence, having itself been fronted, does not intervene between between the fronted INFL and the main verb:

(32) Who came to dinner last night?

Here the *wh-* noun phrase presumably moves to absolute initial position, as English questions normally require, and then INFL moves to COMP. Since the subject does not intervene between INFL and the verb, however, affixation is possible and no *do* will be inserted.[15]

In affirmative declarative sentences like those in (21) above, loss of verb raising to INFL forces the appearance of *do* only in the emphatic cases. In later Middle and in Modern English, one way of insisting on the truth of a sentence is to stress the auxiliary verb, as in (33) and (34):

(33) a. ...whereby máy grow, & dóth growe, in diuerse parties, greate mischef. (244:13:13)
b. God háthe and dóthe shewe for hem gret miracles. (274:8:6)

(34) a. John hás gone to school.
b. John wíll go to school.

This pattern seems to be linked to the development of the special class of Modern English auxiliaries. Unequivocal cases of emphatic *do* do not appear in Middle English until the general periphrastic pattern is well established (Ellegård, 1953), and we have found no evidence of cases like (33) and (34) in Old or early Middle English. If the INFL position does not bear stress in an affirmative declarative sentence, the usual case, then affix hopping is always possible and the use of periphrastic *do* will be at best optional. We will see below, however, that the appearance of *do* in this context is tightly linked to its occurrence in questions and negative sentences. Indeed, the loss of unemphatic affirmative declarative *do* in early Modern English is one of the developments that we hope to illuminate by considering the time course of the change.

It seems clear that the basic syntactic change that transformed the syntax of the INFL or auxiliary verb position between Middle and Modern English was the restriction of verb raising to INFL to auxiliary verbs; and one might, therefore, expect the change in status of the modals to be a reflex or an extension of that change. As we will see below, the loss of main verb V-to-I raising can be dated rather precisely to the middle of the 16th century, the period when the modals become fixed as auxiliaries. The temporal coincidence suggests to us that the modals' change in status is more a result of the syntactic change than a cause of it (*contra* both Lightfoot, 1979 and Roberts, 1985). How the change in the modals might have taken place is indicated by the behavior of the other auxiliaries. Of central importance is the fact that, although *have* and *be* have both auxiliary and main verb uses, they behave syntactically in early Modern English like auxiliaries even when they are main verbs semantically. They always raise to INFL when no auxiliary precedes them, and they appear in non-tensed forms when they are complements to higher auxiliaries. In addition, they never co-occur with periphrastic *do*, even when used as main verbs. This auxiliary-like behavior is entirely carried over into Modern English with the verb *be* and to a certain extent with the verb *have*, as the following examples illustrate:

(35) a. They will not be aware of the danger.
b. They are not aware of the danger.

---

[15] This analysis of the absence of *do* insertion in subject *wh-* questions is not standardly accepted for reasons that go beyond this discussion. Alternative descriptions (e.g., Koopman 1984) are also problematic, however. If the analysis is incorrect, then the absence of *do* insertion in examples like (32) may not provide evidence for the standard account of *do* in questions, but it will provide no evidence against that analysis, which is firmly grounded.

c. Are they aware of the danger?
d. *They did not be aware of the danger.

(36) a. She must not have much money.
b. She hasn't much money. [British]
c. Have you much money? [British]
d. Does she have much money? [Not attested in early Modern English]

The above pattern indicates that Modern English *be*, and early Modern English *have* as well, are somehow single lexical items despite their varied uses. As such, they exhibit the same susceptibility to V-to-I raising both as main verbs and as aspectual auxiliaries. Presumably, they had the same lexical identity in Middle English; but the general applicability of the V-to-I rule to all verbs obscured this fact. When V-to-I raising became restricted to auxiliaries, the lexical identity of these verbs forced them to be categorized as auxiliaries in all their uses and uniformly to continue to exhibit V-to-I raising. Since as aspectual auxiliaries *have* and *be* sometimes raise and sometimes do not, depending on whether or not they are dominated by a higher verb and whether they occur in a finite or a non-finite clause, main verb *have* and *be* also show this variation in where they occur and whether they raise. Given their unitary morphology, the modals must also have counted as single lexical items in Middle English, even though they had somewhat different positional restrictions in their root and in their epistemic uses and in spite of the differences in meaning of these uses.[16] Since only the epistemic modals were semantically auxiliaries, the placement of those modals that counted as auxiliaries in Middle English would have been limited to the INFL position, as in German; only in their root, hence non-auxiliary, uses would the modals appear in non-finite positions or as main verbs. When the restriction of V-to-I raising to auxiliary verbs came in, it forced all modals to behave alike positionally, just as it forced all uses of *have* and *be* to behave alike. Since the epistemic auxiliaries, unlike *have* and *be*, had always been generated under INFL, the root modals also came to be restricted to that position. Thus, we have an explanation for why the root modals lost their ability to function as main verbs and to appear in non-finite positions at the same period in history that V-to-I raising for main verbs was lost.

3.2. The time course of the change.

3.2.1. Given the grammatical analysis presented above, the rise of periphrastic *do* in late Middle English must be a reflex of the loss of V-to-I movement for main verbs. As such, it should be linked to other reflexes of that loss in its temporal development. In this section, we investigate quantitatively the character of this link. Ellegård's extensive study of the rise of periphrastic *do* (Ellegård, 1953) traces its development in a number of contexts. Those which we will initially examine are negative sentences and questions. Figure 6 below, adapted from Ellegård, traces the development and Table 3 gives the data from which the curves in the figure were drawn:

---

[16] One strong piece of evidence that the root modals were not distinct lexical items from the epistemic ones is the fact that they did not develop productively the ability to take *to* infinitive complements. In early Middle English the bare infinitive complement was extremely common, as it had been the predominant form in Old English; but in the course of Middle English, the bare infinitive was generally replaced by the *to* form. This replacement did not take place for the complements of modal verbs, whether root or epistemic. Given the generality of the replacement, the fact that the modals resisted it in both uses argues that the two were tightly linked.

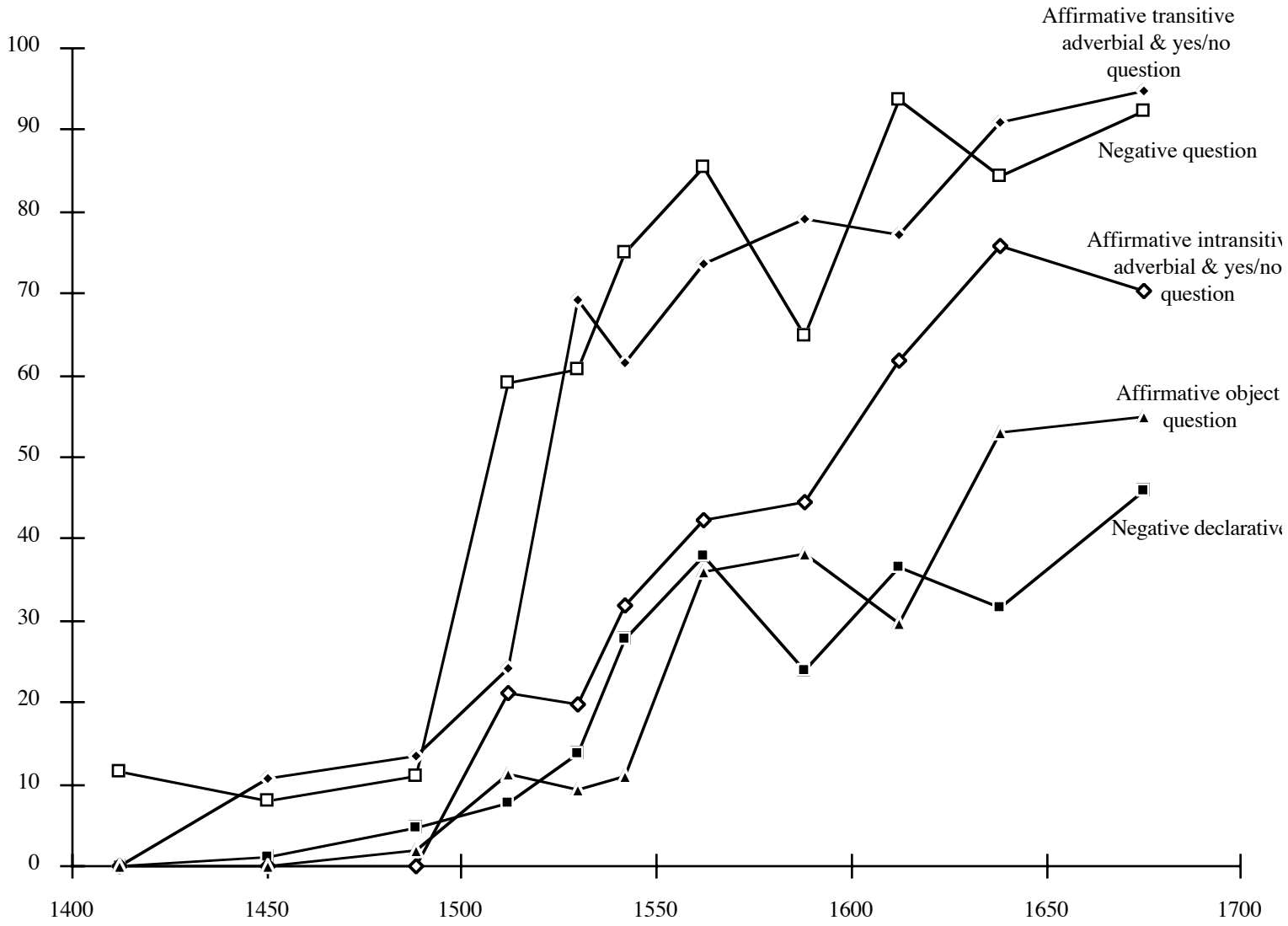


Figure 6: The rise of periphrastic *do* (adapted from Ellegård, 1953).

Period	Date	Negative declaratives		Negative questions		Aff. trans. adv.[17] & yes/no questions		Aff. intrans. adv. & yes/no questions		Affirmative *wh-* object questions	
		% do	total	*% do*	total	*% do*	total	*% do*	total	*% do*	total
Period 1	1400-1425	0	177	11.7	17	0	3	0	7	0	1
Period 2	1425-1475	1.2	903	8.0	25	10.7	56	0	86	0	27
Period 3	1475-1500	4.8	693	11.1	27	13.5	74	0	68	2.0	51
Period 4	1500-1525	7.8	605	59.0	78	24.2	91	21.1	90	11.3	62
Period 5	1525-1535	13.7	651	60.7	56	69.2	26	19.7	76	9.5	63
Period 6	1535-1550	27.9	735	75.0	84	61.5	91	31.9	116	11.0	73
Period 7	1550-1575	38.0	313	85.4	48	73.7	57	42.3	71	36.0	75
Period 8	1575-1600	23.8	629	64.8	128	79.2	173	44.4	205	38.3	120
Period 9	1600-1625	36.7	278	93.7	95	77.3	277	61.9	310	29.8	171
Period 10	1625-1650	31.7	344	84.2	38	90.9	66	75.7	74	53.0	66
Period 11	1650-1700	46.0	274	92.3	52	94.7	76	70.2	131	54.9	51

Table 3: Percentages and totals for the rise of periphrastic *do* (data from Ellegård, 1953).[18]

Figure 6 shows that period 7, from 1550 to 1575, is a point of inflection in the curves of change. Up to then, all of the contexts increase along 'S'-shaped curves; but afterwards negative sentences decline for a time and the affirmative *wh-* object questions remain constant for half a century. Also, as we shall see below, unstressed *do* in affirmative declarative sentences, a feature of Early Modern English which was lost by the eighteenth century, reaches its peak in this period and afterwards declines steadily. It seems plausible to hypothesize that the point of inflection in period 7 corresponds to a major reanalysis of the English auxiliary system. We will defend this hypothesis below; but for the moment let us concentrate on the developments up to that period, which will provide evidence for the character of that reanalysis. As we mentioned in section 1.2 above, if we model 'S'-shaped curves like those in Figure 6 with the logistic function, we can calculate parameters that fix their relationship to one another, most interestingly their rates of increase. To the eye, it seems that the curves in the figure are increasing at different rates, with the higher curves increasing faster. This appearance, however, is caused by the fact that the curves reach the near vertical parts of their 'S's at different times, which logistic curves may do even when their slope parameters are identical. Therefore, to answer the question of whether the curves have the same or different slopes, we must, for each linguistic context, fit the data in Table 3 for periods 1 through 7 to the logistic function[19] and compare the slope and intercept parameters of the fits. When we do this, we obtain the results in Table 4:

---

[17] Adverbial questions are *wh-* questions introduced by *wh-* adverbs like *when*, *where*, *why*, *how*, and others.

[18] The data in this table are taken from Ellegård's Table 7, page 161 and Table 20, page 204. Note that Ellegård groups together adverbial and yes/no questions, having found that these types show almost the same frequency of *do* use in the first half of the change (periods 0 through 7). In the second half of the change, adverbial questions behave like *wh-* object questions but Ellegård continues to group them with yes/no questions.

[19] The regression technique we have used here is a univariate version of the maximum likelihood fit used in the VARBRUL program.

Negative declaratives		Negative questions		Aff. trans. adv. & yes/no questions		Aff. intrans. adv. & yes/no questions		Affirmative *wh*- object questions	
slope	intercept	slope	intercept	slope	intercept	slope	intercept	slope	intercept
3.74	-8.33	3.45	-5.57	3.62	-6.58	3.77	-8.08	4.01	-9.26

Table 4: Slope and intercept parameters of logistic regressions on data in Table 3 (slope is measured in logit units per century).

It is clear by inspection that the slope parameters for the different linguistic contexts are close to one another, and in the categories for which there is the most data, affirmative adverbial & yes/no questions and negative declaratives, they are essentially identical. The range of variation across all five categories is .56, or 15% of the median slope value. If we fit a single best slope to all five contexts, this common slope has a value of 3.70; and under a χ^2 test of significance, the probability of finding, by random fluctuation, deviations from the common slope as large as those in our table is greater than .95 ($\chi^2 = .504$). The results thus support the hypothesis that the slopes of the curves are underlyingly the same and that the observed differences among them are random fluctuations. The intercept values in Table 4 are based on fixing the zero point in time (t = 0) at 1350 A.D. This date is a reasonable one since the rate of use of periphrastic *do* in Ellegård's prose sample, the data that we are reanalyzing, is very close to zero at that period (below .01%). Mathematically, the choice of zero point is arbitrary since the logistic curve extends from $-\infty$ to $+\infty$, and it does not affect the regression estimates of the slope of the curves. However, if we can estimate the beginning point of the change empirically and we set t = 0 at that point, we can directly test our constant rate hypothesis against one of the alternatives discussed in section 1.2; namely, that the rate of increase differs by context while the initial frequency of use of *do* is the same in all contexts. If t=0 at 1350, we find that the hypothesis of same initial frequency is rejected with a probability less than .025 ($\chi^2 = 11.41$).[20]

3.2.2. We are now in a position to tackle the basic question posed jointly by our grammatical analysis of the change in the English auxiliary system and our hypothesis on the relationship of changes across contexts: Can the rise of in use of periphrastic *do* be related quantitatively to any other reflex of the loss of V to I raising? Under our hypothesis, we would expect any other such reflex to change at the same rate as the use of *do*. If we find such a result, our constant rate hypothesis will be considerably strengthened since the changing forms will not exhibit any superficial relationship to one another. Fortunately, Ellegård's study contains quantitative data on one reflex of the loss of V to I movement other than the rise of periphrastic *do*; namely, the changing position of unstressed adverbs. Although he does not relate this change grammatically to periphrastic *do*, Ellegård notices that contemporaneously with the rise of *do* , the position of unstressed sentence adverbs shifts from post-verbal to preverbal position in sentences with tensed main verbs. Thus, as we noted above in section 3.1, in Middle English weak adverbs like *never* and *always* ordinarily occur after the tensed main verb while in Modern English they occur before it, the contrast being illustrated in (37) and (38):

(37) Quene Ester looked never with swich an eye.

---

[20] The result is significant for any choice of t=0 later than 1300. As we move the zero point back in time, the differences among the intercepts, measured in percentages, quite naturally decrease as they all decline toward zero. As this happens, the differences among them become less significant statistically. The point at which the statistical significance of the differences crosses the conventional .05 level is 1300 ($\chi^2 = 7.87$, $.1 > p > .05$). I would like to thank Saul Sternberg for a helpful discussion of the statistical issues involved here.

(Chaucer, *Merchant's Tale*, line 1744)

(38) Queen Esther never looked with such an eye.

This difference is, in the analysis we have adopted, due to the respective application and non-application of V to I raising in the Middle English and Modern English examples. Our constant rate hypothesis predicts, therefore, that the replacement of the order in (37) by that in (38) should occur at the same rate as the rise in use of periphrastic *do*. In the discussion which follows, we will demonstrate that this prediction holds.

The data given by Ellegård pertain to the position of the adverb *never*, and we give them here in Table 5:

Period	Date	*do never* V	*never* V	V *never*
Period 2	1425-1475	3	52	99
Period 3	1475-1500	4	80	102
Period 4	1500-1525	1	80	28
Period 5	1525-1535	3	151	16
Period 6	1535-1550	14	125	13
Period 7	1550-1575	9	71	8
Period 8	1575-1600	6	152	5

Table 5: The position of *never* with respect to the main verb (Ellegård, 1953:184).

As the table shows, the analysis of the position of adverbs in late Middle and early Modern English is somewhat more complex than the analysis of periphrastic *do*, in questions and negative sentences. First of all, there are three variants instead of two. When the tensed main verb raises to INFL, we find the order V-*never*. In the absence of V to I raising, two further variants are possible. Either the tense/agreement affix moves onto the verb,[21] yielding the order *never*-V, or an unemphatic periphrastic *do* appears in the INFL position and carries the affix. Since we are interested in the rate of loss of V to I raising, however, we can reduce this three way alternation to a two way alternation by grouping the two non-raising alternants together. In other words, it is the frequency of the V-*never* alternant against the sum of the two others which we must track over time. A further problem arises, however, when we realize that the frequency of the order *never*-V is surprisingly high even in the first period in Table 5. At a time when the use of *do* is still rare, this order occurs in one third of cases. While it is possible that the appearance of a sentence adverb could have such a strong disfavoring effect on V to I raising, this seems unlikely. Rather, the figure is misleading because the word order in question can arise in two different ways. If the position of the adverb is between INFL and VP (see the diagram in (25) above), then the *never*-V order indicates that affix hopping has occurred. However, there is another position that the adverb can occur which gives the word order a different interpretation. In late Middle English, as in Modern English, weak sentence adverbs can occur immediately preceding INFL as well as immediately following it. The example below illustrates this possibility:

---

[21] When a sentence adverb occurs between INFL and VP, affixes must move across it when they attach to a verb which remains *in situ* as the head of VP. If we follow current thinking in transformational grammar, the fact that this movement is possible indicates that the adverb is not the head of a phrasal projection but rather an adjunct to VP.

(39) For many are that never kane halde the ordyre of lufe... (Rolle, "The Bee and the Stork," lines 20-21 [excerpt in Mossé (1952)])

The appearance of *never* before the auxiliary verb gives clear evidence that this pre-INFL positioning of the adverb is possible.[22] When no auxiliary is present, it is not possible to distinguish the case of INFL-*never*-V order with affix hopping from that of *never*-INFL-V order. Only the former cases are relevant to the the loss of V to I raising since in the latter cases the raising of V to I has no effect on word order. We must conclude that the data in Table 5 are in themselves insufficient to give an accurate measure of the change we are considering. In order to use them, we must estimate the fraction of surface order *never*-V cases that are reflexes of a *never*-INFL-V structure and subtract that fraction from the cases in the table. We have calculated such an estimate based on the instances of *never* in two concordances to 14th century literature, Tatlock and Kennedy's Chaucer concordance (1927) and Kottler and Markman's concordance to five late Middle English poems (1966). All of these works date from the late 14th century, when the appearance of periphrastic *do* is no more than sporadic and when we can assume that the frequency of V to I raising is still essentially categorical. Table 6 gives the figures for the two concordances:

Source	*never* - INFL	INFL - *never*	% pre-INFL *never*
Chaucer	5	29	14.7
Five Poems	21	101	17.2

Table 6: Frequency of pre-INFL *never* before the rise of affix hopping.

The percentages of pre-INFL *never* for our two samples are remarkably close, suggesting that the simple average of the two figures (16.0 %) will give a reliable estimate of the population frequency. If we further assume that this percentage is constant through the period of change, we can estimate, via logistic regression, the rate of loss of V to I raising in the data of Table 5. The assumption that the percentage is constant is an empirical one which might be false. It seems reasonable, however, given that the pre-INFL position remains a possible but marked option in Modern English.[23] In any case, our results are only moderately sensitive to the exact level of pre-INFL *never*. Figure 7 below shows the decline of the V-*never* order with both Ellegård's figures and our adjustment:

---

[22] It is not clear from the data we have thus far examined whether this pre- INFL position is one adjoined to the INFL node or to one of its phrasal projections. For our purposes this distinction is not important since we are attempting to estimate the frequency of pre-INFL *never* that does not reflect affix hopping in the data as Ellegård himself coded them. Since he did not distinguish among various grammatical configurations that would have led to this order, we also cannot, if we want our figures to be comparable to his. Clearly, more work reanalyzing Ellegård's data is needed.

[23] In the Brown corpus of American English, out 327 cases of *never* immediately before or after an auxiliary verb, 34, or 10.4%, are pre-INFL. This estimate is not very different from the estimates we have for late Middle English and strengthens our assumption that the rate is constant in time. I would like to thank Beatrice Santorini of the Penn Treebank Database Project for calculating this estimate for me.

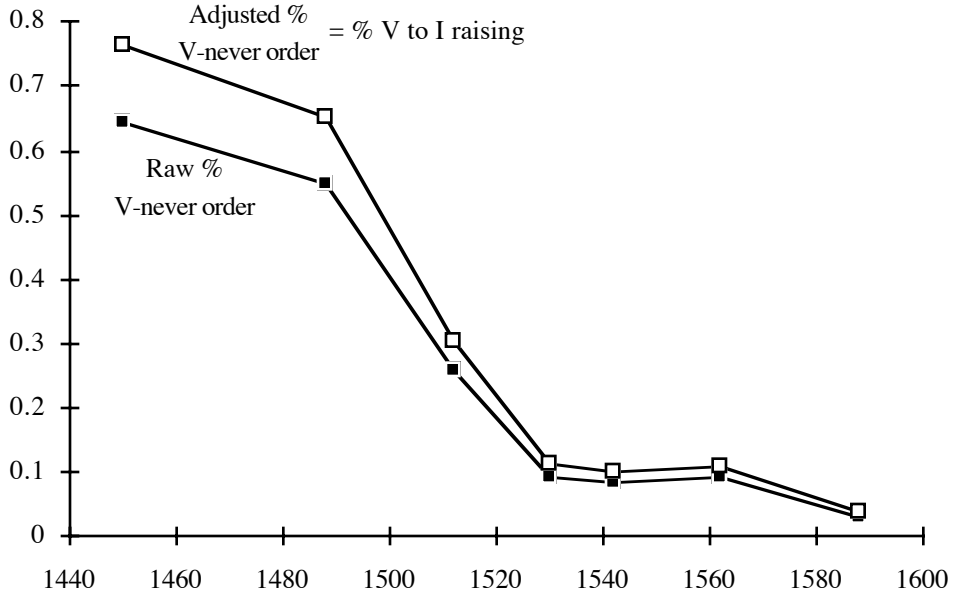


Figure 7: The decline of V to I raising in sentences with *never* .

Table 7 below gives the intercept and slope values for the loss of V to I raising based on Ellegård's *never* data with five different values for the percentage of pre-INFL *never*, our estimate and two values on either side of it:

% pre-INFL *never* =5.0		% pre-INFL *never* =10.0		% pre-INFL *never* =16.0		% pre-INFL *never* =20.0		% pre-INFL *never* =25.0	
slope	intercept	slope	intercept	slope	intercept	slope	intercept	slope	intercept
-3.38	4.43	-3.53	4.76	-3.76	5.37	-3.97	5.78	-4.32	6.45

Table 7: Slope and intercept parameters of logistic regressions on data in Table 5 for five values of percentage pre-INFL *never*.

The slopes here are negative and the intercepts positive because we are now looking at the change as the loss of V to I raising, which simply reverses the signs of the parameters. It is clear that our best estimate yields exactly the expected result. The slope of the curve is essentially identical to that for the rise of *do* in questions and negative sentences ($\chi^2 = .028$, $p > .8$). Even when we allow our estimate of the percentage of pre-INFL *never* to vary by about 10% above and below the estimate based on the concordances, the range of slope values is not much larger than that among the various contexts for *do*, as given in Table 4. Only the highest of these values (the rightmost column in Table 7) is significantly different from the best fit common slope of the contexts in Table 4. Thus, we have here substantial evidence that all contexts reflecting the loss of V to I raising change at the same rate.

3.2.3. Let us turn now to a final context for the appearance of periphrastic *do*, the unemphatic affirmative declarative case. As we mentioned above in section 3.1, late Middle and early Modern

English allow the auxiliary *do* to appear in ordinary affirmative declarative sentences like (21), repeated here as (40):

(40) a. They worschipped the sonne whanne he dede arise. (78:327:8)
b. When he dyd se[e] that Crist schold be dede.... (167:188:2)
c. Me thinke I doe heare a good manerly Begger at the doore...(346:5:17)

The interest of this context is that, unlike the others we have examined, the appearance of *do* here does not correspond directly to the loss of V to I raising. Since, in this case, affix hopping is not blocked by the appearance of negation or by movement of INFL to COMP, the failure of V to I raising does not force the use of *do*. Indeed, the frequency of periphrastic *do* never goes above 10% in simple affirmative declarative sentences. Figure 8 and Table 8 trace this development:

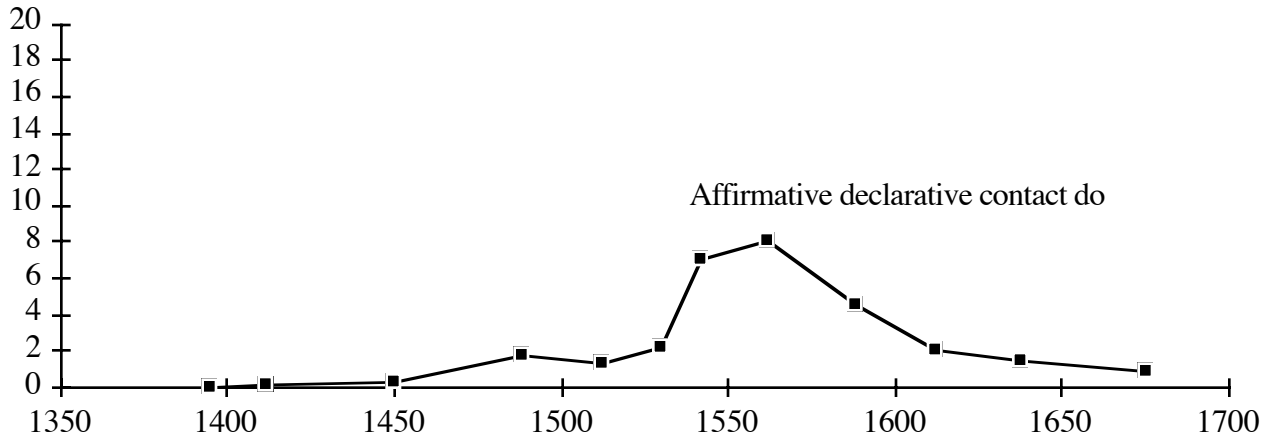


Figure 8: Percent *do* in unemphatic affirmative declarative sentences (contact cases only[24]).

[24] These data are drawn from Ellegård (1953) like all of our other data. They are limited to cases of what Ellegård calls 'contact' *do*; that is, cases where *do* immediately precedes the main verb. The other cases of affirmative declarative *do* discussed by Ellegård are those where an adverb intervenes between *do* and the main verb and those where the subject has inverted with *do*. These should also be analyzed but cannot be, because Ellegård does not give breakdowns of the frequency of the non-*do* alternant by subtype. Because the contact cases are 94% of the total affirmative declarative *do* cases in the beginning periods of the change, we have estimated the number of affirmative declarative non-*do* contact cases at 94% of the total of non-*do* declaratives reported by Ellegård. Since the number of non-contact cases is small, the results we report do not change in any significant way if all affirmative declarative subtypes are lumped together, the only alternative that Ellegård's categorization of his data makes available to us.

Period	Date	% *do*	total[25]
Period 0	1390-1400	0.014	42300
Period 1	1400-1425	0.23	4324
Period 2	1425-1475	0.27	42770
Period 3	1475-1500	1.78	56024
Period 4	1500-1525	1.37	26884
Period 5	1525-1535	2.27	17672
Period 6	1535-1550	7.05	18048
Period 7	1550-1575	8.13	13724
Period 8	1575-1600	4.59	16920
Period 9	1600-1625	2.07	7426
Period 10	1625-1650	1.43	6768
Period 11	1650-1700	0.92	7426

Table 8: Percentages and totals for the rise of periphrastic *do* in contact affirmative declaratives (Ellegård, 1953).

These data exhibit three interesting features. First of all, although the rate of *do* use is much lower in affirmative declarative sentences than in any other context, its use in these sentences begins no later in time. Because affirmative declarative sentences are by so much the most frequent kind of sentence, the number of potential instances of *do* is much greater in that context than in any other. Hence, although the frequency of *do* use is low in this context, the actual number of instances found in texts is greater than for any other context; and this is true from the earliest appearance of periphrastic *do*. Thus, due to the accidental fact that the context least favoring to *do* use is the most common one, we have evidence against the sequential actuation scenario described in section 1.2 and in support of simultaneous unequal actuation of the change. Secondly, it is striking that the rate of use of *do* in the affirmative declarative context rises along with the other contexts until the period (period 7) when other contexts stop rising together, which suggests that the context is tied to the others even though it is not a simple reflex of the loss of V to I raising. Thirdly, the curve seems to level off in the middle of the sixteenth century at a frequency of *do* of just under 10%. This leveling off is not consistent with the logistic model, under which the slope of the curve should increase continuously until a frequency of 50% *do* is reached, and it indicates that a somewhat more complex mathematical model may be needed for the affirmative declarative context than for the others.

[25] These totals were estimated by Ellegård on the basis of exhaustive sampling of the occurrences of *do* and small random sample of the non-*do* cases. Ellegård proceeded in this fashion because, as is obvious, the number of affirmative declarative sentences is so large. Counting each of them would not appreciably improve the accuracy of the estimate over that based on a random sample. Interested readers are directed to Ellegård's own discussion (Ellegård 1953: 157-159).

If we attempt to fit a logistic curve to the data in Table 8, the results are as in Table 9:

Affirmative declarative contact *do*	
slope	intercept
2.82	-8.32

Table 9: Slope and intercept parameters of logistic regressions on data in Table 8.

The slope estimate for this case is significantly different from the 3.70 value that gives the best fit to the questions and negative sentences ($\chi^2 = 9.64$, $p<.01$). This result confirms that affirmative declarative *do* should not simply be assimilated to the other contexts.[26] We are left then with the question of why the use of *do* in this context increases regularly up to period 7 and whether these data contradict the constant rate hypothesis. In addition, we would like to know why the use of affirmative declarative contact *do* declines after period 7, just when the other contexts begin to move independently. While definitive answers to these questions will require further research, there are interesting indications in Ellegård's data. One plausible assumption that would lead to the observed rise in frequency of affirmative declarative *do* through the middle of the sixteenth century is the following: Suppose that the relative frequencies of *do* use and affix hopping were constant; that is, that, throughout this period, the set of sentences produced without V to I raising exhibited a constant ratio of *do* use to affix hopping. Then, given such a constant proportion and given that the frequency of sentences without V to I raising increases with time, the frequency of affirmative declarative *do* would necessarily also increase. Note that for any assumed proportion of *do* use to affix hopping, the frequency of V to I raising for each time period could be estimated. From these estimates, the rate of loss of V to I raising could then be estimated as the slope of a logistic regression and compared to the slope obtained for the other contexts. Table 10 below shows the relationship between the proportion of *do* use assumed and the parameters of the logistic regression, as the proportion moves from .05 to .20:

Proportion of *do* =.05		Proportion of *do* =.10		Proportion of *do* =.15		Proportion of *do* =.20	
slope	intercept	slope	intercept	slope	intercept	slope	intercept
4.84	-7.59	3.79	-7.14	3.39	-7.16	3.21	-7.21

Table 10: Regression parameters for affirmative declarative *do* for five values of the conditional probability of *do*, given that V to I raising has not applied.

While we do not have any accurate way of measuring the proportion of *do* use in the absence of V to I raising, we can reasonably assume that it is above .081, since that is the frequency of use in

---

[26] The result differs from that reported in Kroch (1989), where affirmative declarative *do* is reported to increase at the same rate as the other contexts. In the earlier work fewer contexts were analyzed and the method used for estimating the logistic parameters was less accurate than the ones we currently have available.

period 7, at the highest point of the curve for the affirmative declarative contact cases. If at that point the language no longer exhibits any V to I raising, then the observed frequency will directly estimate the proportion. If, on the other hand, there is still a small amount of V to I raising in that period, then the observed frequency will underestimate the proportion. In the *never* data discussed earlier, V to I raising still applies in 11% of cases in period 7, giving an estimate of the proportion of *do* of .091. As Table 11 shows, assuming either .081 or .091 as the proportion of *do* yields a slope very close to the 3.7 that we would expect on the basis of our other results.

Proportion of *do* =.081		Proportion of *do* =.091	
slope	intercept	slope	intercept
4.03	-7.18	3.87	-7.21

Table 11: Regression parameters for affirmative declarative *do* with conditional probability of *do* estimated from other data.

The χ^2 measures of the difference of fit between the slopes above and 3.7 are very low: .50 and .16 respectively, with $p>.40$ and $p>.65$. Once again, though here depending on an auxiliary assumption, we find support in Ellegård's data for the constant rate hypothesis and for a grammatical analysis that treats the development of Modern English auxiliary verb syntax as a reflex of the loss of V to I raising.

3.3. Grammatical reanalysis as reflected in the frequency data on *do*.

We now conclude our discussion of periphrastic *do* with a look at the problem of why the contexts of Table 3 evolve independently of one another after period 7. A glance at Figure 6 shows that *do* use in negative declarative sentences falls briefly in period 8 and does not resume its regular upward movement until period 11. Questions, on the other hand, show only small deviations from a continued monotonic increase in the use of *do*. Most strikingly, of course, Figure 8 shows the use of affirmative declarative *do* declining regularly in the later periods. Under the analysis we have suggested for this last case, the inflection point in the curve is at the point when V to I raising is lost from the grammar. At that point, the affirmative declarative context comes to allow only two options, affix hopping and periphrastic *do*, instead of the three allowed in the preceding period. Presumably, these two options begin to compete with one another just when V to I raising disappears; and affix hopping eventually wins out over the use of *do*. Before the loss of V to I raising, we have assumed, the two options exist in fixed relative proportions and so are not in competition. Why the competition should begin only when V to I raising is lost is at present unclear and represents an important unresolved issue in our interpretation of the history of the affirmative declarative context.

Another issue, but one on which we have relevant data, is that of how the loss of V to I raising goes to completion. One possibility is that the change goes to completion successively in its various contexts; V to I raising is lost first from ordinary affirmative declaratives, next from questions and last from negatives. Such a successive completion scenario would correspond, for the end of a change, to the successive actuation scenario at the outset of a change, against which we gave some evidence (see sections 1.2 and 3.2). We have somewhat more detailed evidence that the successive completion scenario is also incorrect. When we measure the rate of change in the

frequency of *do* across contexts after period 7, we find the following results for the contexts we have been tracking:

Affirmative declarative contact S's		Negative declaratives		Negative questions		Aff. trans. adv. & yes/no questions		Aff. intrans. adv. & yes/no questions		Affirmative *wh* object questions	
slope	intercept	slope	intercept	slope	intercept	slope	intercept	slope	intercept	slope	intercept
-2.29	-2.45	.497	-.947	1.42	.870	1.36	.830	1.30	-.329	.743	-.810

Table 12: Slope and intercept parameters of logistic regressions on *do* data from period 7 to period 11 (intercepts are set at 1562).

Clearly, the rate of change in the use of affirmative declarative *do* is not the same as the rate of change in other contexts. It is not only significantly different from those contexts, but is significantly different from zero ($\chi^2 = 62.5$, $p<.0001$), reflecting the consistent decline in its use over the century and a half after the mid-sixteenth century. On our analysis, this result is expected once V to I raising has been lost, whether it is lost everywhere at once or first in the affirmative declarative cases. When we look at the negative declarative and question contexts, however, the results are more informative. In those contexts, the use of *do* continues to increase, but the rate of increase slows dramatically. In every context the rate is less than half of what it had been (with a significance level of $p<.001$ in all cases). This pattern does not fit easily into a successive completion scenario, since under that scenario it is unclear why, when one context goes to completion, the rate of change in other contexts should be affected. On the other hand, if the loss of V to I raising is simultaneous in all contexts, then those for which surface reflexes of the raising remain after the point of loss must have been reanalyzed grammatically. As a result of the reanalysis, new grammatical options come to be in competition and the rate of evolution of this new competition will, in general, differ from the rate of change among the different grammatical options that were competing before reanalysis. In other words, the simultaneous completion scenario requires that in negative sentences and questions, the non-*do* forms change their grammatical analysis at the point where V to I raising is lost. Forms that were reflexes of V to I raising (fronted tensed main verbs and tensed main verbs followed by *not*) must now be generated by other mechanisms. The plausibility of the scenario thus depends on whether there is evidence for such reanalysis. In fact, such evidence does exist, though it is certainly not conclusive at the present state of our understanding.

Not only does the use of *do* increase more slowly in questions and negative declaratives after period 7 than before, but it also no longer increases at the same rate in all of these contexts. The question contexts do continue to move together, as is evident by inspection for all of the question contexts except object questions, which have a somewhat lower slope that the others. The difference between the object questions and others is, however, not statistically significant ($\chi^2 = 2.99$, $p>.4$) and is probably due to the fact that the curve for object questions is somewhat irregular in the periods immediately after period 7, as can be seen from Figure 6. The negative declarative context, on the other hand, increases at a significantly slower rate than the question contexts ($\chi^2 =14.0$, $p>.01$) and shows considerable irregularity of shape. This divergence will only be consistent with our constant rate hypothesis if the grammatical reanalysis at period 7 splits the grammatically unitary process of change of the early periods into three independently evolving processes in the affirmative declarative, negative declarative and question contexts, respectively. In the affirmative declarative context it is clear what forms are in competition: Affix hopping is competing with *do* support of the tense and agreement affixes in INFL. In the negative declarative and question contexts, on the other hand, the nature of the grammatical alternatives is harder to determine. The issue, of course, is what grammatical analysis to assign to the non-*do* negatives and questions once V to I raising is no longer available. Recent work on the grammar of INFL and its affixes has

interesting implications in this area, which are worth some comment, though it is clear that much further work will be needed.

Pollock (1989) argues that *do* support is obligatory in Modern English negative sentences because English *not* is the head of a phrasal projection, NegP, which blocks affix hopping. In this, *not* is to be distinguished from *never*, which, as a VP adjoined adverb, has no such blocking effect. This difference in grammatical status is reflected directly in the following contrast, familiar from our earlier discussion:

(41) a. John never sings rock and roll.
b. *John does never sing rock and roll.

(42) a. John does not sing rock and roll.
b. *John not sings rock and roll.

In early Middle English, *not* was probably not the head of NegP but rather an adverbial adjunct like *never*. The head of NegP would have been the preverbal negative particle *ne*, which appeared before INFL, rather than after it. As the sentences in (43) below indicate, *not* was a second negative element, which behaved much like French *pas*:

(43) a. For hire seolf ne kepte heo (she) nawt of the world.(mid thirteenth century, *St. Katherine*, folio 11:32)
b. ...ye ne ought nat as for hir deeth yourself to destroye. (late fourteenth century, Chaucer, *Tale of Melibee*, line 84).

The particle *ne* disappeared from English as a sentence negator at about the time that *do* began to spread. From the middle of the 14th century, *not* begins to appear alone (Mossé, 1952:112), and this usage takes over rapidly. Thus, though *ne* is still widespread in Chaucer, it is already uncommon in the Wycliffite sermons and other late fourteenth and early fifteenth century texts. For some time after *not* appears without *ne*, it is probably still not the head of NegP but remains a VP adjunct. The NegP may continue to exist with an empty head in place of the old *ne*, as apparently happens in certain Romance dialects (Zanuttini, 1989). Some evidence for this analysis of early *not* is that it sometimes appears before a tensed main verb. This order, illustrated in (44) below, is rare; but in the later 16th and the 17th centuries it is a live option, while in earlier Middle and in later Modern English it is extremely rare (Visser, 1969:1532-33):

(44) a. ...they deafe men's eares, but not edify. (1593, Thomas Nashe, *Christ's Teares*, *Works*, II:123)
b. ...he that filches from me my good name robs me of that which not enriches him. (1604, Shakespeare, *Othello*, III:iii:161)
c. Safe on this ground we not fear today to tempt your laughter by our rustic play. (1637, Ben Jonson, *Sad Shepherd*, Prologue:37)

Visser says that when *not* appears before the tensed verb, it is appearing in the position of earlier *ne* (in what we would now call the pre-INFL position); but this is unlikely. Evidence against this analysis is that the order *not* followed by tensed verb does not occur when the tensed verb is an auxiliary. Visser gives 59 examples of *not* before a tensed verb and in all of them, the tensed verb is a main verb. There are no cases with modals or aspectual auxiliaries and not even any with the main verbs *be* or *have*. Since the verb *to be* is by so much the most frequent verb in texts, the absence of it and other auxiliaries from Visser's list of examples is telling. What Visser's examples actually indicate is that in the transition period between Middle and Modern English *not* did not block affix hopping. Its surface position was always between INFL and VP, where it was left adjoined to VP; and, as with Modern English *never*, affixes could move across it to attach to the main verb when the latter did not raise. The temporal coincidence of the free occurrence of this order with the loss of V

to I raising thus provides additional support for the grammatical analysis the change in English auxiliaries we adopted in section 3.1. It also gives a hint as to the nature of the reanalysis that might have taken place when V to I raising was lost from the grammar of English. When raising was lost, negative sentences where *not* followed the main verb could have been reanalyzed as resulting from enclisis of *not* onto the end of the following tensed main verb. This enclisis would have been optional, though heavily favored. When it applied, it would have changed the superficial order of verb and negation from *not*-V to V-*not*; and when it did not apply, examples like (44) would have resulted. We know that *not* does become an enclitic on tensed auxiliaries in this period and it is not a very bold speculation to suggest that enclisis onto a tensed main verb that has not raised to INFL should also be possible, though more work will be needed to confirm or falsify this analysis. If we accept it provisionally, we have an explanation for why the use of *do* in negative sentences does not rise in the period immediately after 1575. After the reanalysis, the use of *do* in negative sentences becomes as redundant as it is in affirmative declaratives. Since *not* is not the head of NegP and so does not block affix hopping, the use of *do* has little motivation. For this reason, it does not advance, even falling in the period immediately after the loss of V to I raising. Later, toward the middle of the seventeenth century, another reanalysis occurs which forces the use of *do* in negatives; namely, the shift of *not* from VP adjunct to head of a post-INFL NegP. Wherever this change occurs, the use of *do* becomes obligatory, as *not* in its new guise as a phrasal head does block affix hopping. The gradual rise of *do* in negatives (*do* is not categorical in negative sentences until the nineteenth century (Rydén, 1979)), then reflects the competition between the "VP adjunct" and the "head of NegP" analyses for *not*.

In questions, the loss of V to I raising should also lead to a grammatical reanalysis of those sentences without periphrastic *do*. The most obvious reanalysis would be to a grammar that allowed direct fronting of tensed main verbs from their underlying head of VP position to COMP, in place of movement through INFL. This direct V to COMP movement would then compete with the use of periphrastic *do* and its fronting from INFL to COMP. In time, the latter would win out. Under this scenario, we face the question of how the Middle English grammar would have blocked direct V to COMP movement, and, in particular, why it became possible just when V to I raising was lost. A recent paper by Platzack and Holmberg (1990) suggests, at least tentatively, an answer to these questions and contributes to the plausibility of postulating the needed reanalysis. The authors argue that, among the VO (verb-object order) Germanic languages, direct V to COMP movement is possible in exactly those which do not have V to I raising and instead use affix hopping to inflect tensed verbs. V to I raising, they find, occurs in languages with subject-verb agreement (Icelandic, Middle English, Old Scandinavian, and Yiddish) while affix hopping occurs in those lacking agreement (modern mainland Scandinavian). The presence of the agreement morpheme in INFL blocks direct movement of a verb from VP to COMP. When agreement is absent, INFL neither provides a landing site for the verb nor blocks movement to COMP. Under Platzack and Holmberg's analysis, Middle English behaves as expected: it has rich subject-verb agreement and V to I raising. Modern English, however, is a problematic case. It retains only radically impoverished agreement morphology and V to I raising of the two verbs *have* and *be*; and unlike other languages with agreement, including Middle English, it does not allow sentences with oblique or with empty subjects, two features that Platzack and Holmberg link to a full agreement morpheme in INFL.[27] Nevertheless, Modern English does not allow direct movement of tensed main verbs to COMP. It thus appears to be an intermediate system, whose agreement morphology is rich enough to block V to COMP movement but insufficient for the licensing of empty and oblique subjects. Perhaps the intermediate status of English reflects the fortuitous existence of periphrastic *do* at that point in the history of Middle English when the loss of much agreement inflection weakened the motivation for the V to I raising analysis, at least of simple sentences where

[27] A similar problem arises in the case of French, which clearly has V to I raising but lacks the other properties Platzack and Holmberg associate with strong INFL. To cope with these cases, they argue that, in non-verb second languages V to I raising is possible without strong agreement. Whether this approach can be maintained is unclear; but in any case it is beyond the scope of our brief discussion.

V to I movement does not alter word order. As V to I raising came to be restricted to auxiliary verbs, the existence of auxiliary *do* would have provided a derivation for main verb questions without V to COMP movement. At the point where V to I raising was lost, those questions in which main verbs were still fronted to COMP would have been reanalyzed as exhibiting V to COMP movement, a possible analysis due to the weakness of English agreement inflection. This analysis would have required treating remaining subject-verb agreement (the second and third person singular) as entirely superficial, with no syntactic significance.[28] Presumably, the treatment of agreement as non-syntactic, which might have led to its eventual loss (as in Scandinavian), was in competition with a continuing analysis of agreement as syntactic in the ordinary way, which availability of *do* continued to make possible. It is not at all clear, however, why that analysis of agreement and the concomitant obligatory use of *do* in questions should have won out, as it obviously did. Further work is clearly needed here, especially given the fact the second person singular form (*thou* V+*st*) disappears, thus further weakening agreement, just as the use of *do* becomes categorical.

4. Conclusion.

The idea that language change proceeds context by context, with new forms appearing first in a narrowly restricted context and spreading to others only later, has been widely accepted. It has seemed obvious that the ordering of contexts in the spread of a change reflected the linguistic forces causing the change. Just as the discovery that a given social group most uses a certain innovation is taken to show that the innovating form originated in that group, so the finding that a given context is most favorable to the use of an innovation is taken to show that the innovation is an accommodation to the linguistic functionality of that context. Similarly, just as the process by which other social groups come to use an innovation is assumed to be active or passive imitation of the innovating group, so the spread of a change from context to context is assumed to be due to analogy, also a kind of imitation, where speakers imitate their usage in one context in the linguistic options they choose in another. One consequence that flows from this conception of change is that quantitative studies can be taken to provide direct evidence as to the causation of change, either regarding the origin of innovations or the functional pressures that favor their advance. The contexts which favor an innovative form functionally are also those in which it arises or advances most rapidly. Hence, examination of those contexts and contrasting them with disfavoring contexts should reveal the causes of the change and/or the forces responsible for its advance. A recent example of this reasoning is work by Stein (1986) on the case of periphrastic *do* itself. In a quantitative analysis of a corpus drawn from Shakespeare's plays and some other near-contemporary work, Stein finds that the rate of use of *do* in questions correlates with the complexity of the consonant clusters produced when verbal endings are affixed to given stems. When the use of *do* would eliminate a phonotactically marked sequence of consonants at the end of the verb, its appearance is favored. For example, in the past tense of weak verbs, addition of the second person singular ending *-st* produces the marked cluster 'dst', as in *showed - showedst*. When the verb stem ends in a

---

[28] Whether such an approach is correct is beyond the scope of this paper, but it is certainly conceivable. The case of agreement with postposed subjects in Modern Standard English existential sentences provides a clear example of agreement which is superficial in the required sense. Thus, it may seem from sentences like (i) that postposed subjects agree with their verbs in the ordinary way; but as the examples in (ii) show, the agreement does not respect the syntactic constituency of the subject in the way that agreement for preverbal subjects does:

(i) a. There was a box on the table.
b. There were/*was two boxes on the table.

(ii) a. There was/?were a box and a lamp on the table.
b. A box and a lamp *was/were on the corner.

consonant, the cluster produced is even more marked, as in *stopped - stoppedst*. Thus, *do* use is favored in the past tense of weak verbs when the ending is second person, even more so when the verb stem is consonant final. Having found this effect, Stein argues that he has found the cause of the rise of periphrastic *do*; namely, the elimination of phonotactically marked clusters. In cases where the use of *do* has no such effect, as in *goes he* versus *does he go*, Stein says that *do* comes in by "generalisation" from its use in the other cases. The only evidence that Stein gives to support his analysis is that of the correlation between *do* use and phonotactic factors. He must, therefore, be assuming that discovery of the contexts that favor the use of *do* suffices to find the cause of its rise. Like Bailey and others, he does not defend the assumption with evidence.

Our results cut against the assumption that Bailey, Stein and others have made and that we also accepted in earlier work (Kroch *et al*., 1982; Kroch, 1982). They demonstrate instead that change proceeds at the same rate in all contexts, and that, as far as one can tell, disfavoring contexts acquire new forms no later than favoring ones, though at lower initial frequencies. If this conclusion is correct, then the pattern of favoring and disfavoring contexts does not reflect the forces pushing the change forward. Rather, it reflects functional effects, discourse and processing, on the choices speakers make among the alternatives available to them in the language as they know it; and the strength of these effects remains constant as the change proceeds. In the case of Stein's work, this would mean that the phonotactic effects on *do* use, while real, would be constant during throughout the change, and would coexist with other, also constant, factor effects, like those Ellegård investigated - negation, transitivity, and so forth. None of these effects would have any privileged causal status. Indeed, from the evidence that Stein presents, it does seem that the phonotactic effects show the expected constancy. Thus, his data show *wh-* questions always using less *do* than yes/no questions, even where the phonotactic contexts are the same; and more significantly, the size of the effect of various favoring and disfavoring phonotactic contexts is largely constant across time. Stein himself notes that

> ...corpora [generally] move upwards in all categorial values [i.e., in all contexts - A.K.] in a parallel fashion: that is, in their drift towards 100% periphrasis they increase their share of periphrastic tokens in all categories, but they keep a fairly constant distance in all subcategories to diachronically earlier corpora (Stein, 1986:137-38).

This result of Stein's is, of course, just what we would expect; but for him it is striking, and requires a special explanation. Thus, he sees the phonotactic effects as responsible for initiating the rise in periphrastic *do* from an initial low frequency; but having done so, the effects cease to operate and spread by analogy takes over, pushing all contexts up together. In the absence of supporting evidence, such a change in the causal efficacy of the phonotactic effect seems implausible.

In conclusion, the work reported here strengthens certain affiliations of historical quantitative studies and renders others problematic. We have been able to show, with statistical methods, the controlling effect of abstract grammatical analyses on patterns in usage data. When surface forms change, the new usage reflects a change in the underlying grammar that licenses the forms; and incremental linguistic change seems often to reflect competition among alternative licensing principles for entire grammatical subsystems. Further work on historical change promises to extend the evidence linking patterns of change to grammars in competition, allowing us to understand changes better from the perspective of linguistic theory and eventually, perhaps, to refine grammatical analyses on the basis of the predictions they make about the patterning of usage in change.[29] On the other hand, if the constant rate hypothesis is correct, attempts to understand the internal linguistic causes of language change from the study of usage data will have to be redirected. Not only will it no longer be possible to induce causes from contextual effects in the way that many

---

[29] Once the principle that contexts change together when they are surface reflexes of a single grammatical competition becomes firmly established, it may be possible, on occasion, to choose among grammars proposed on the basis of synchronic analysis by the predictions they make as to which contexts should change together.

have done, but it will also be necessary to look for causes of change at more abstract levels of structure. When replacements involve entire grammatical subsystems, as in the case of the loss of the verb-second constraint in French or the loss of V to I raising in English, the causes of the replacement must lie at the level of the entire subsystems in competition and not at the level of the specific linguistic contexts in which the competition is observed. Thus, since V to I raising in English is lost in all finite clauses with tensed main verbs and at the same rate, there must be a factor or factors which globally favor this loss. Here we differ from the discussion in Kroch (1989). In that paper, we proposed differential misunderstanding as a possible mechanism driving syntactic change, the idea being that when two linguistic alternatives were in competition, one might be more often misunderstood than the other. If such differential misunderstanding obtained, then under reasonable assumptions (see Kroch, 1989 for details), the less often misunderstood form would in time replace its competitor; and the plot of frequency of the advancing form against time would follow the logistic function. In our view, this model remains a plausible one; but we no longer believe that the data we presented in support of it can actually provide evidence in its favor. We argued that the tendency of transitive questions and transitive negative declaratives to show more *do* use than intransitives (see Table 3 above and Ellegård, 1953:192-207) showed that the advantage of *do* use was in its allowing the main verb of a transitive sentence to remain adjacent to its direct object, thereby simplifying parsing. From our current perspective, this explanation cannot be correct in the form in which we gave it. Since the difference between the transitive and intransitive contexts is constant across time, our 1989 discussion is identical in logic to the one by Stein that we rejected above. Unfortunately, the evidence that would be needed to actually show that some linguistic factor had a causal effect in advancing a change under the differential misunderstanding model, or under any functional model as far as we can see, would seem to be of a kind unavailable to historians. It would have to be a demonstration that the causal factor indeed was correlated with differential misunderstanding and that the differential misunderstanding effect was global. Although an advancing form might have an advantage only in certain contexts, the effect of its use on the rate of misunderstanding summed across all contexts would have to favor it in order for the effect to obtain. The fact that contexts tied together grammatically do not evolve independently means that functional effects can drive change only to the extent that their average effect on usage is to favor one form over its competitor. Demonstrating the existence of functional effects would seem, therefore, to be a problem best left to studies of contemporary usage, where experimental methods can be applied. If solid results can be obtained from such experiments, they will almost certainly shed new light on problems of historical change. Once again, the present would be used to explain the past.

5. References.

Adams, Marianne Patalino. 1987a. From Old French to the Theory of Pro-Drop. NLLT 5 : 1-32.

Adams, Marianne Palatino. 1987b. *Old French, Null Subjects, and Verb Second Phenomena*. Ph. D. dissertation, UCLA.

Aldrich, John and Forrest Nelson. 1984. *Linear Probability, Logit, and Probit Models*. *Quantitative Applications in the Social Sciences*. London: Sage Publications.

Altmann, G. *et al*. 1983. A Law of Change in Language. In *Historical Linguistics*. Edited by B. Brainard. 104-115. Bochum, West Germany: Studienverlag Dr. N. Brockmeyer.

Bailey, Charles-James. 1973. *Variation and Linguistic Theory*. Washington: Center for Applied Linguistics.

Bock, J. Kathryn and Anthony Kroch. 1989. The Isolability of Syntactic Processing. In *Linguistic Structure in Language Processing*. Edited by G. Carlson and M. Tannenhaus. Boston: Kluwer.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, Noam. 1986. *Barriers*. Cambridge: MIT Press.

Denison, David. 1985. The Origins of Periphrastic Do: Ellegård and Visser Reconsidered. In *Papers from the 4th International Conference on English Historical Linguistics*. Edited by Roger Eaton *et al*. Amsterdam: Benjamins. 45-60.

Ellegård, Alvar. 1953. *The Auxiliary do: The Establishment and Regulation of Its Use in English*. Edited by Frank Behre. *Gothenburg Studies in English*. Stockholm: Almqvist and Wiksell.

Emonds, Joe. 1978. The Complex V - V' in French. LI 9 (2) : 151-175.

Engblom, W. 1938. *On the Origin and Early Development of the Auxiliary Do*. Lund: C. W. K. Gleerup.

Fontaine, Carmen. 1985. *Application de Methodes Quantitatives en Diachronie: L'Inversion du Sujet en Français*. M.A. thesis, Université du Québec à Montréal.

Hirshbühler, Paul and Anne Rochette, eds. 1988. *Aspects de la Syntaxe Historique du Français*. Association Québécoise de Linguistique.

Hirshbühler, Paul and Marie-Odile Junker. 1988. Remarques sur les Sujets Nul en Subordonnées en Ancien et en Moyen Français. In *Aspects de la Syntaxe Historique du Français*. Edited by Paul and Anne Rochette Hirshbühler. 63-84. Association Québécoise de Linguistique.

Joshi, A.K. 1985. Processing of Sentences with Intrasentential Code-switching. In *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*. Edited by D. Dowty, L. Karttunen, and A. Zwicky. 190-205. Cambridge: Cambridge University Press.

Kemenade, Ans van. 1989. The Diffusion and Implementation of the Category Modal in Middle English. Paper presented at International Conference on Historical Linguistics IX, Rutgers.

Koopman, Hilda. 1984. *The Syntax of Verbs*. Dordrecht: Foris.

Kottler, Barnet and Alan M. Markman. 1966. *A Concordance to Five Middle English Poems*. Pittsburgh: University of Pittsburgh Press.

Kroch, Anthony, Susan Pintzuk, and John Myhill. 1982. Understanding 'do'. In *Papers from the 18th Regional Meeting of the Chicago Linguistics Society*. Edited by K. Tuite *et al*. Chicago: Chicago Linguistics Society.

Kroch, Anthony. 1982. Grammatical Theory and the Quantitative Study of Syntactic Change. Paper presented at NWAVE 11. Georgetown University:

Kroch, Anthony, Beatrice Santorini, and Caroline Heycock. 1987. Bare Infinitives and External Arguments. In *Proceedings of the 18th Annual Meeting of the North Eastern Linguistics Society*. Edited by James Blevins and Juli Carter. Amherst, MA: Department of Linguistics, University of Massachusetts.

Kroch, Anthony. 1989. Function and Grammar in the History of English: Periphrastic 'do'. In *Language Change and Variation*. Edited by Ralph Fasold. Amsterdam: Benjamins.

Labov, William. 1975a. The Use of the Present to Explain the Past." In *Proceedings of the 11th International Congress of Linguists*. Edited by L. Heilmann. 825-851. Bologna: Il Molino.

Labov, William. 1975b. *What is a Linguistic Fact?* Lisse, Netherlands: Peter de Ridder Press.

Labov, William. 1982. Building on Empirical Foundations. In *Perspectives on Historical Linguistics*. Edited by W. Lehmann and Y. Malkiel. 17-82. Amsterdam: Benjamins.

Lightfoot, David. 1979. *Principles of Diachronic Syntax*. New York: Cambridge University Press.

Lightfoot, David. 1988. Syntactic Change. In *Linguistics: The Cambridge Survey. Volume 1*. Edited by Frederick Newmeyer. 303-323. New York: Cambridge University Press.

Lightfoot, David. 1989. The Child's Trigger Experience: Degree-0 Learnability. BBS 12 (2) : 321-334; commentary 334-375.

Mossé, Fernand. 1952. *A Handbook of Middle English*. Translated by James A. Walker. Baltimore: The Johns Hopkins Press.

Noble, Shawn. 1985. To have and have got. Paper presented at NWAVE 14. Georgetown University:

Oliveira e Silva, Giselle. 1982. *Estudo da Regularidade na Variação dos Possessivos no Português do Rio de Janeiro*. Ph. D. dissertation, Universidade Federal do Rio de Janeiro.

Osgood, Charles, and Thomas Sebeok. 1954. Psycholinguistics: A Survey of Theory and Research Problems. *J. Abn. and Social Psych.* 49 (4, part 2) : 1-203.

Piotrovskaja, A. and R. Piotrovskij. 1974. Matematiceskie Modeli Diachronii i Tekstoobrazovanija. In Statistika Reci i Avtomaticeskij Analiz Teksta. 361-400. Leningrad: Nauka.

Platzack, Christer and Anders Holmberg. 1990. The Role of AGR and Finiteness in Some European VO Languages. University of Lund: unpublished manuscript.

Pollock, Jean-Yves. 1989. Verb Movement, Universal Grammar and the Structure of IP. LI 20 (3) : 365-424.

Priestley. 1955. Reprise Constructions in French. *Archivum Linguisticum* 7 : 1-28.

Prince, E.F. 1984. Topicalization and Left-Dislocation: a Functional Analysis. In *Discourses in Reading and Linguistics. Annals of the New York Academy of Sciences*. Edited by S.J. White and V. Teller. vol. 433. 213-25.

Prince, E.F. 1988. Discourse Analysis: a Part of the Study of Linguistic Competence. In *Linguistics: the Cambridge Survey. Vol. 2. Linguistic Theory: Extensions and Implications*. Edited by F. Newmeyer. 164-82. New York: Cambridge University Press.

Roberts, Ian. 1985. Agreement Parameters and the Development of the English Modal Auxiliaries. NLLT 3 (1) : 21-58.

Rousseau, Pascale and David Sankoff. 1978. Advances in Variable Rule Methodology. In *Linguistic Variation: Models and Methods*. Edited by David Sankoff. 57-69. New York: Academic Press.

Rydén, M. 1979. *An Introduction to the Historical Study of English Syntax*. Stockholm: Almqvist & Wiksell.

Sankoff, D. and Poplack, S. 1981. A Formal Grammar for Code-switching. *Papers in Linguistics* 14 : 3-46.

Santorini, Beatrice. 1989. *The Generalization of the Verb-Second Constraint in the History of Yiddish*. Ph. D. dissertation, University of Pennsylvania.

Saussure, Ferdinand de. 1966. *Course in General Linguistics*. New York: McGraw-Hill Book Company.

Spiess, Eliot. 1989. *Genes in Populations*. New York: John Wiley and Sons.

Steele, Susan *et al*. 1981. *The Encyclopedia of AUX. A Study of Cross-Linguistic Equivalence*. Cambridge: MIT Press.

Stein, Dieter. 1986. Syntactic Variation and Change: The Case of do in Questions in Early Modern English. *Folia Linguistica Historica* VII (1) : 121-149.

Tatlock, John and Arthur Kennedy. 1927. *A Concordance to the Complete Works of Geoffrey Chaucer*. Washington: The Carnegie Institution.

Tukey, John. 1977. *Exploratory Data Analysis*. Reading, MA: Addison-Wesley Publishing Company.

Vance, Barbara. 1988. L'évolution de Pro-drop en Français Médiéval. In *Aspects de la Syntaxe Historique du Français*. Edited by Paul and Anne Rochette Hirshbühler. 85-109. Association Québécoise de Linguistique.

Visser, Frederick Th. 1963-1973. *An Historical Syntax of the English Language*. Leiden: Brill.

Warner, Anthony. 1982. *Complementation in Middle English and the Methodology of Historical Syntax*. University Park: The Pennsylvania State University Press.

Warner, Anthony. 1983. Review of D. Lightfoot, *Principles of Diachronic Syntax*. *J.of Ling*. 19 : 187-209.

Weinreich, Uriel, W. Labov and M. I. Herzog. 1968. Empirical Foundations for a Theory of Language Change. In *Directions for Historical Linguistics: A Symposium*. Edited by W. P. Lehmann. 95-195. Austin: University of Texas Press.

Woolford, E. 1983. Bilingual Code Switching and Syntactic Theory. LI 14(3) : 520-36.

Zanuttini, Raffaella. 1989. "The Position of Negative Markers in Romance Languages." Paper presented at the GLOW Colloquium, Utrecht.