

FUNCTION AND GRAMMAR IN THE HISTORY OF ENGLISH:
PERIPHRASTIC *DO*

ANTHONY S. KROCH
University of Pennsylvania

Abstract.

The introduction of periphrastic *do* into English questions and negative sentences is sufficiently well documented that quantitative research on the phenomenon is feasible. It then becomes possible to observe the course of the change in considerable detail. The study suggests that an ongoing syntactic change can continue while the grammar remains fixed--a change that shows up in differences in relative frequencies of competing forms, all of which are allowed by the grammar. Only when one form displaces the others entirely will there be a reorganization of the grammar. When historical change is understood in this way, it is possible to locate the points at which grammatical reorganization takes place.

1. Introduction.

1.1 In a remarkable early attempt to apply quantitative methods to the study of syntactic change, A. Ellegård (1953) tabulated the relative frequency of periphrastic *do* from its first systematic use in the thirteenth century to the early eighteenth century, when it became obligatory as a support for tense in questions and negative sentences. At the beginning of this period, questions and negatives are formed by rules that place the subject and the particle *not* in that order, after the first verb in the sentence. As the examples in (1) and (2) show, these rules do not differentiate between main verbs and auxiliaries (i.e., the modals and perfect *have*):¹

- (1) a. Wherefor then serveth the latter...? (350 132/21)
b. Wylte thou see by some example that thys is trueth that I nowe saye? (320 408/33)
c. Alack, how long have I prolonged these ancient years and hoar hairs most unhappy...? (344 81/18)
- (2) a. Fly fro company of them that lovyth not honour & trouthe. (312 10/16)

This is an offprint from:

RALPH W. FASOLD and DEBORAH SCHIFFRIN (eds)
LANGUAGE CHANGE AND VARIATION

John Benjamins Publishing Co.
Amsterdam/Philadelphia
1989

(Published as Vol. 52 of the series
CURRENT ISSUES IN LINGUISTIC THEORY)
ISSN 0304-0763 / ISBN 90 272 3546 5

© Copyright 1989 - John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm or any other means, without written permission from the publisher.

- b. Thro whiche labour lytelle men schalle not be inducede oonly to doctrine but also grete men schalle be provocate to(287I 15/10)
 c. The nowble and grete Constantyne made in hit a chirche off Seynte Sepulcre, whiche hathe not suffrede iniury un to this tyme of enmyes of the felthe....(287I 111/13)

The birth of periphrastic *do* in the thirteenth century allows questions and negative sentences with only a main verb, like (1a) and (2A), to be formulated in a new way, as in (3):

- (3) a. Why then doth my Rosalynde grieve at the frown of Torismond, who by offering her a prejudice proffers her a greater pleasure? (353 33/10)
 b. I do not allow but abhorre incontynence in (315 50/11)

At some point in the fourteenth century, these new forms begin to increase in frequency and eventually they come to replace the old--or, as Ellegård calls them, the 'simple'--forms. In the modern language, the use of the *do* form is, of course, categorical in almost all environments. Only where the main verb of a sentence is *be*--or, in some dialects, *have*--does the simple form remain possible. This paper is part of a project currently underway to reanalyze this change, using theoretical advances in syntax and in quantitative linguistics to refine and extend the results of Ellegård's pioneering work (see also Kroch, Pintzuk, and Myhill 1982).²

For his study, Ellegård collected from prose texts,³ more than 10,000 tokens of the alternating *do* and simple forms, comparing their relative frequencies in several syntactic environments and grouping the texts by manuscript date or date of first publication into 12 historical periods. The results of this work are summarized in Table 1 and displayed graphically in Figure 1, which are adapted from his book.

As Figure 1 shows, the replacement of the simple form by the *do* form is a complex affair that extends over many generations. Indeed, Ellegård's figures underestimate the length of time involved, for, as Rydén 1979 points out, the change is not really complete until the nineteenth century.⁴

1.2 The fact that the gradual character of this change is so well documented raises anew the classic difficulty that linguistic theory has had in dealing with

PERIOD	NEG. QUES.		AFF. QUES.		NEG. DECL.		AFF. DECL.	
	#	%	#	%	#	%	#	%
1390-1400	0	00.0	0	00.0	0	00.0	0	00.0
1400-1425	2	11.8	0	00.0	0	00.0	11	00.2
1425-1475	2	08.0	6	04.2	11	01.2	121	00.3
1475-1500	3	11.1	10	07.0	33	04.8	1059	01.8
1500-1525	46	59.0	41	22.7	47	07.8	396	01.4
1525-1535	34	60.7	33	32.4	89	13.7	494	02.6
1535-1550	63	75.0	93	44.9	205	27.9	1564	08.1
1550-1575	41	85.4	72	56.3	119	38.0	1360	09.3
1575-1600	83	64.8	228	60.3	150	23.8	1142	06.3
1600-1625	89	93.7	406	69.2	102	36.7	240	03.0
1625-1650	32	84.2	116	82.9	109	31.7	212	02.9
1650-1700	48	92.3	164	79.2	126	46.0	140	01.8

Table 1. The frequency of *do* by environment. [Ellegård 1953:166]

drift (Sapir 1921, Lakoff 1972, Malkiel 1981). Much current work in historical syntax (e.g., Lightfoot 1979), influenced by the success of syntactic theories based on immediate constituent analysis, assumes that change proceeds by structural reanalysis. This approach entails that change is abrupt, since syntactic structures are topological entities not amenable to incremental modification. Any apparently gradient historical data must be generated by underlying discrete grammatical shifts. One can either postulate a chain of small discrete changes in the underlying grammar (Hausmann 1974)⁵ or appeal to external factors like dialect mixture. Where it can be shown that a historical development proceeds by a sequence of saltations, the hypothesis of a chain of reanalyses is plausible. Where the usage of different authors differs categorically, so that variation exists within the community but not within the usage of individuals, dialect mixture may be appropriate as an explanation.

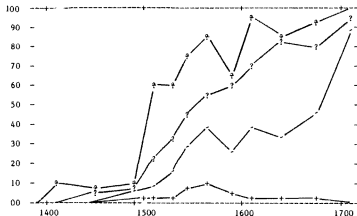


Figure 1. Percent *do* in affirmative and negative questions (? and ?) and affirmative and negative declaratives (+ and -) [Ellegård 1953:162]

In the case of the rise of periphrastic *do*, however, neither of these moves is convincing. The data collected by Ellegård allow us to see that the change proceeds as a gradual increase in the use of the *do* form in each of its linguistic environments rather than by a series of categorical changes in which the *do* form replaces the simple form, environment by environment. Because Ellegård's

collection includes substantial amounts of data from a number of individual authors, moreover, we can see that there is a large amount of variation within the usage of individual authors, who, as far as random variation allows us to judge the matter, seem to share a single pattern of variable use of both new and old forms. In this circumstance, to invoke dialect mixture would be meaningless, since the mixture would be within the individual, and the fact that the mix of forms used is consistent across the community at any one time, but changes in a uniform direction over time, would remain unexplained.

In this paper I argue, using the case of periphrastic *do* as a concrete illustration, that syntactic change can often be better explained in psycholinguistic terms than in purely grammatical terms. My argument appeals to the hypothesis, plausibly though not conclusively demonstrated (Bock 1982), that within the constraints of the grammar of their language, speakers tend, all other things being equal, to construct sentences so as to conserve scarce psycholinguistic resources by minimizing processing complexity.⁶

I bring this assumption to bear on the problem of syntactic change by noting that two competing forms may differ in complexity, so that one form is favored in usage. Because constraints on processing complexity are gradient rather than all-or-nothing in character, the relative frequency of the two forms in usage data reflects the degree of difference in their complexity. Either a change in the mix of forms available (through borrowing and other processes) or a change elsewhere in the syntax of the language (e.g., in the default word order) can alter the relative complexity of forms, thereby triggering change as the usage frequencies accommodate to the new circumstances. In the most extreme cases, the difference in complexity of the two forms will be great enough to drive one form out of the language entirely.

We will expect the accommodation to new circumstances to be gradual, however, both because the changes on which it depends may be gradual and because the relative frequency of use of competing or alternating forms in a given speech community seems to be part of the community's linguistic norms and so learned in the course of language acquisition. This fact gives the frequencies a certain arbitrary or conventional character and means

that they cannot respond instantaneously, but only in a mediated way, to psycholinguistic pressures. This last point is, of course, controversial, and I return to it later in my exposition. One result I hope to establish, at least for the case of periphrastic *do*, is that making the assumption that frequencies of use are part of community linguistic norms--or some assumption with equivalent effect--is justified by the explanatory load that it can be made to bear.

The approach to syntactic change that I have outlined sharply restricts the role of grammatical reanalysis in the historical process. It is only at the endpoints of changes, when disfavored forms go out of use entirely or become conscious archaisms used for literary effect, that we would postulate the occurrence of a grammatical reorganization. During the course of a change the grammar remains fixed while the relative frequencies of the competing forms slowly change under pressure from the processing system. The competing forms may be expected to occur in a number of syntactic contexts, and the advantage that one form has over the other may vary from one context to another. In this circumstance one would expect usage data tabulated by context to reflect this variation, and in the case of periphrastic *do* we find the data showing such patterning. Furthermore, the amount of this variation and the constraints on its expression turn out to tell much about the process of change.

Let me conclude this sketch of my approach to syntactic change by anticipating what is perhaps the most interesting result of the investigation. I have found--and later hope to demonstrate--that once a historical change is analyzed as a gradual change in the relative frequencies of competing forms, it becomes possible to locate the points at which grammatical reorganization occurs. It is even possible, I believe, to obtain indications of the nature of the formal reanalysis by comparing the forms that pattern together in frequency changes before and after the reorganization.

2. A sketch of the history of periphrastic *do*.

According to Ellegård, who here follows widespread scholarly opinion (but see Hausmann 1974), periphrastic *do* develops out of an earlier causative use, analogous to

the causative use of *faire* in French. Some scholars have claimed that periphrastic *do* originates from the Old English pro-verb *do* or some other use, but Ellegård uses quantitative arguments to defend the causative origin theory. He points out that the causative *do* occurred both with and without an overt complement subject, as a comparison of (4) and (5) illustrates.

- (4) Sche dede hym etyn & drynkyn & comfortyd hym...
(193 239/7)

'She made him eat and drink and comforted him.'

- (5) He did make haules and chambres riche. (81

64/2)

'He had beautiful halls and rooms built.'

In the environment with overt complement subject, this construction is clearly differentiated from an ordinary, noncausative sentence; but in the environment without overt complement subject, the distinction is less certain. The empty subject was interpreted as unspecified, just as it is in French; but this interpretation was also available for an ordinary simple sentence, since a sentence like 'John built a house' could (and still can) mean either that John constructed a house with his own hands or that he had it built. The consequence of this overlap is that there are many occasions when there is no difference in meaning between a sentence containing causative *do* without overt complement subject and one without *do* in which the verb is interpreted causatively. Compare sentence (6) with (7):

- (6) In token that he had myght, a kastle he did
reweye. (81 96/24)

'As a sign of his strength, he [the king] had a castle built.'

- (7) ...the tours bette he doun. (81 97/22)

'He [the king] knocked the towers down.'

Ellegård calls examples like (6) equivocal in order to contrast them with cases like (4), which must be interpreted as causative. Because a sentence containing causative *do* without overt complement subject may be synonymous with a simple sentence whose verb is interpreted causatively, the verbal complex in (6) is open to a permutation of meaning from causative *do* + non-causative main verb to periphrastic *do* + causative main verb. Such a reinterpretation preserves the meaning of

the sentence as a whole while changing the respective semantic contributions of the two verbal elements.

This potential reinterpretation, Ellegård argues, becomes actual when unambiguously causative *do* (i.e., *do* with overt complement subject) is replaced *make*, for this replacement removes the best evidence that *do* carries a causative meaning. The development occurs first in the dialects of the southwest, where *make* is from early on the favored causative in cases with overt complement subject. It then spreads eastward to the dialects that are the most direct antecedents to modern standard English, and eventually to the north.

While the explanation that Ellegård proposes for the origin of periphrastic *do* is not novel (see Engblom 1938), he goes further than previous historical linguists in using quantitative data to prove his case. In particular, he shows that as causative *do* with overt complement subject is first challenged and then replaced by *make*, undoubted cases of periphrastic *do* first appear and then rise in frequency relative to the equivocal cases--those that are equally interpretable as causative or as periphrastic. The progress of this change is illustrated in Table 2.

VERSE	periphrastic	0	15	253
	equivocal	4	164	124
	with compl. subj.	66	58	16
PROSE	periphrastic	0	0	2
	equivocal	0	15	91
	with compl. subj.	8	18	36

Table 2. Causative, equivocal and periphrastic *do* in Eastern texts. [Ellegård 1953:45]

As Table 2 suggests, causative *do* with overt complement subject more or less disappears by some time in the fifteenth century, and from then on one can conclude that all the equivocal cases are periphrastic.

Once periphrastic *do* appears, its history has two main phases. During the first, when it still coexists with causative *do*, it occurs with very low frequency, primarily in poetry, where it functions to allow placement of the verb, in the infinitive form, at the end of the line for purposes of rhyme and meter. As far as one can tell, given its low rate of use, periphrastic *do* occurs at this period in all the environments in which it later becomes important. Thus, in the following famous example from Chaucer *do* occurs first in a question and then in a declarative.

- (8) His yonge sone, that three yeer was of age
 Un-to him seyde, fader, why do ye wepe?
 Whan wol the gayler bring our potage
 Is ther no morsel breed that ye do kepe?
 ("Monk's Tale," lines 441-444)

Since the vast majority of sentences are affirmative declaratives, however, most of the examples of periphrastic *do* from this period are also affirmative declaratives.

The second phase in the history of *do* is the one represented in Figure 1. It begins at the end of the fourteenth century, when causative *do* is disappearing, and lasts until the eighteenth century. In this period, the overall frequency of the *do* form rises rapidly, and its rate of use stratifies increasingly by environment. Until 1560, the frequency of *do* in all environments increases monotonically, though seemingly more rapidly in the more favored contexts. After 1560, the frequency of *do* in affirmative declarative sentences declines steadily until, by 1700, the modern prohibition against the use of *do* in this environment is essentially established.⁸ Negative sentences, both declarative and interrogative, decline briefly with the affirmative declaratives and then rise to join the affirmative questions in categorically requiring the use of *do*.

3. Processing factors in the rise of periphrastic *do*.

3.1 Ellegård's analysis shows clearly that the replacement of the simple verb by the *do* form in questions and negative sentences is consequent on the general shift in word order that marks the transition from Middle to Modern English. This transition, which has been dis-

cussed from various perspectives by many scholars (for instance, see Fries 1969, Haiman 1974, Sweet 1899, and Canale 1978), changed English from an SOV language that exemplified in modified form the Germanic verb second principle to one with strict SVO order, in which no constituent may appear between a verb and its object.

The change in basic word order from SOV to SVO may date from as early as the twelfth century (Canale 1978); it also appears to trigger a number of further syntactic changes. The rise of periphrastic *do* is one such change. Another is the progressive loss of inversion between subject and verb in sentences that are introduced by an adverb or topicalized noun phrase (Jacobsson 1951, Schmidt 1980). This Middle English allowed, although it did not require, subject-verb inversion in sentences like the following:

- (9) Thus departed the queene in the company of the
sayd syr John lorde Beaumont, who ryght joiously
dyd conducte her to Valencyennes.... (309 30/29)
(10) That tour founded kyng Nembroth. (255 25/13)

As Jacobsson 1951 shows, this inversion becomes less and less common between the fifteenth and eighteenth centuries. According to his statistics, inversion after a sentence-initial adverb occurred in 44% of cases in the fifteenth century, but by the seventeenth century the rate had fallen to 7%. Ellegård adds to this finding the observation that inversion is much more disfavored in transitive sentences, where the inverted subject would separate the verb from its object, than in intransitive sentences (p. 188 ff.).

A third change consequent on the shift to SVO word order and roughly contemporary with the other two is the shift in position of a class of adverbs of time and modality, including *never*, *always*, *often*, and others, out of postverbal position. Before 1400, these adverbs were most likely to appear after the tensed verb, whether that verb was the main verb or an auxiliary; beginning in the fifteenth century, however, there is an increasing tendency for the adverb to appear immediately before a tensed main verb, although in sentences with auxiliaries, its position does not change. This shift may also be related to the strengthening prohibition on the appearance

of phrases between the verb and its object, since these adverbs would have appeared there in transitive sentences.⁹

The contribution made by periphrastic *do* to the establishment of the new principles of word order becomes clear from a comparison of *do* and simple forms like those in (11) and (12).

- (11) a. How like you this sonnet? (353 84/5)
b. Assuredly there is nothing that can bee
perfectly gotte, either through labour, or
through learning, if man grounde not his
doinges altogether upon Nature. (338 50/8)
(12) a. Doest thou ask me, Saladyne, for they cates?
(353 12/8)
b. Whereupon he did not make the wife upon the
same clay, whereof he made man.... (338
40/8)

As these examples show, in simple questions the inverted subject was placed between the main verb and its complement, and in simple negative the modal adverb *not* was placed after the tensed main verb. The use of the *do* forms, by contrast, maintains the adjacency of verb and object under question inversion and removes *not* from the position after the main verb. This latter function allows *not*, which as an enclitic, must immediately follow the tensed verb,¹⁰ to participate in the word order shift that the other adverbs of its class are undergoing.

In a recent reanalysis of the data on questions in Ellegård's study (Kroch, Pintzuk, and Myhill 1982), we found that the evidence for the conclusion that the rise of periphrastic *do* is part of the word order shift between Middle and Modern English is even stronger than Ellegård's own tabulations suggest. Thus, he found that transitive sentences were more likely to contain *do* than intransitives, as would be expected if the rise of periphrastic *do* were part of the word order change; we discovered several additional effects pointing to the same conclusion. For example, in questions, the frequency of *do* was affected by whether the subject, which appears between the verb and its object when the verb is in the simple form, was a pronoun or a full NP. In the former case, the *do* form was less likely, presumably because the

pronoun was a less salient intervening element between verb and object than was a full NP. Indeed, since subject pronouns were clitics when they appeared in immediate postverbal position (Mossé 1952), one might want to argue that no independent constituent intervenes between the verb and object in a sentence like (13).

- (13) O thou Rhamnusia, o thou goddesse of indygnation, whiche (revengist the upon prowde folkes) howe playste thou the stepdame with me? (326 149/16)

A second support for the link between the rise of periphrastic *do* and the general word order shift is the fact that *do* is no more likely to occur with verbs taking sentential complements than it is with intransitives. This result is as expected because there is no requirement of adjacency between verb and complement in this case. Thus, while adverbs cannot occur between a verb and its direct object in Modern English, they still occur freely between verb and sentential complement. The contrast is illustrated in (14).

- (14) a. *John saw clearly the light.
b. John saw clearly that the light was on.

Third, we found that transitive question in which the object was a pronoun and the subject a nonpronominal NP were the most likely environment to show the *do* form. As the examples in (15) show, simple inversion in this environment forces the subject NP to come after the object pronoun, which obligatorily cliticizes onto the fronted verb.

- (15) a. Wherefore lighteth me the sonne? (304 25/5)
b. And wherefore doth the earth sustain me? (304 25/24)

Simple inversions like (15a) provide the only case in the language in which the subject of a verb follows the object. Thus, a word-order-based theory of the loss of inversion of subject and main verb predicts that their use would be heavily disfavored once an alternative way of forming questions came into existence.

Finally, we found that an accumulation of clitic forms after the verb also favored the use of *do*, so that examples like (16) are quite heavily disfavored compared to other quantifiers with pronoun subjects.

- (16) a. Know ye me nat? (243 975/6)
b. Toke ye hym in the quenys chamber? (243 1174/6)

This effect is not directly related to the word order change, but it shows how the use of *do* eliminates an awkward effect of simple subject-verb inversion--namely, the piling up of unstressed syllables.¹¹

3.2 Let us take as demonstrated the hypothesis that the rise of periphrastic *do* is part of the contemporaneous general word order shift that English undergoes. We can now proceed to see how a processing-oriented theory of change can elucidate the connection between the two developments. Because we still know so little about how people actually process sentences it is not yet possible to prove the connection and specify it exactly. What we can do is to see how certain general assumptions about the constraints under which sentence processing operates can provide the basis for an explanation of the change. I will therefore sketch an account based on current psycholinguistic literature and explore its consequences for our understanding of the historical material. To the extent that the account elucidates the material, we have reason to accept it--at least provisionally--and to pursue further research along these lines.

Some recent attempts to model sentences processing (Frazier and Fodor 1978, Carroll 1981) have argued that sentence comprehension requires two separate processing modules. One is a parsing routine that assigns constituent structure to sentences. The other is a segmentation procedure, logically--although perhaps not temporally--prior to the first, which breaks up the incoming string into clause-sized units.

Carroll 1981 proposes that the segmentation procedure is actually a lexical structure analyzer that isolates units comprising predicates and their arguments for further analysis. Because these lexical structures must be isolated without reference to the hierarchical grouping of words into phrases, Carroll suggest that the

predicate and its arguments are identified by local signals of form class and grammatical function. More precisely, he sees the segmentation module working by identifying the verb in a clause through morphological criteria like tense, aspect, and agreement marking, and then grouping with it surrounding NPs according to the argument structure assigned to it in the lexicon and the casemarking on the NPs. In an analytic language like English, the role of morphological marks is partly taken over by function words, like the prepositions and infinitival *to*, which appear adjacent to the argument NPs.

Not surprisingly for an area of research in which so little is yet known, the idea of a two-module parser is controversial. Indeed, much current opinion would argue that including a segmentation module in the parser is both unnecessary and unworkable (Wanner 1980). On the other hand, Carroll's idea that local cues are essential to the identification of verbal arguments (or at least of the boundaries of argument phrases) is considered plausible even by researchers who reject his overall views (Hindle, personal communication).

This situation, while less clear-cut than one might like, is adequate to our purposes, for it is the notion that local cues are important to the identification of verbal arguments on which this discussion depends. That hypothesis of itself can, I believe, explain why English develops a constraint against the appearance of nonparenthetical material between verb and direct object, as SVO word order comes to predominate.¹²

The explanation is a simple one. Since by the period in question, English has lost its nominal case system entirely, prepositions, which are always adjacent to their objects, have become the local signal for complement NPs. Only the direct object lacks this signal, and adjacency to the verb comes to serve as the local cue to its argument status. Whenever a phrase appears between the verb and object, the identification of lexical structure becomes more difficult. Whenever syntactic options are available that remove the intervening material, they will be favored in usage and by an amount proportional to the degree to which they simplify analysis. For example, we expect that in a question, an inverted subject pronoun, because it is a clitic, will have less effect on the recoverability of the argument status of a direct ob-

ject NP than would an inverted full NP subject. Therefore, we predict the constraint described in section 3.1, whereby questions with pronoun subjects favor the use of *do* less than those with full NP subject.

Similarly, because *not* is also a clitic, we expect it to appear more frequently in postverbal position than independent negative adverbs like *never* or *seldom*—an expectation also confirmed by the data (Ellegård 1953:184). Because *never*, *seldom*, and other adverbs are not syntactic dependents, they may appear before the main verb without the introduction of *do*. Thus, the gradual change in their position from postverbal to preverbal is grammatically distinct from the change in position of *not*. Nonetheless, it is a response to the same word order change that drives the rise of *do*. The fact that the tendency toward preverbal placement of independent adverbs is stronger than the tendency favoring the use of *do* in sentences with *not* argues for our account of the forces involved.

The explanation I am proposing unifies the processing constraint underlying the rise of *do* with 'heaviness' phenomena still active in the modern language. Thus, the one construction in which a constituent ordinarily falls between the verb and direct object is the double object dative, which alternates with prepositional datives introduced by *to* or *for*. As is well known, the choice between these two options in usage is greatly influenced by the heaviness of the two argument NPs. The same heaviness effect governs the relative positioning of particle and object NP in verb particle constructions. In this case we again find that pronouns (here the direct object pronoun) have special constraints on their position in the verb phrase due to their clitic status.

Heaviness also comes into play when the direct object NP is long and complex and the verb takes an additional argument. Thus, while sentence (17) is unacceptable, (18a) is more natural than (18b) because the direct object NP contains a relative clause.

(17) *I carved with a knife the turkey

(18) a. I carved with a knife the turkey that had been put in front of me.

b. I carved the turkey that had been put in front of me with a knife.

It is widely assumed that heavy NP shift has processing motivations, and under the assumptions I have been making, the motivation is that of reducing the load on memory of the process of searching for the arguments of the verb. By shifting the heavy argument to the end of the VP, the parser is able to identify those arguments more quickly and presumably more efficiently. The cost of this shift is, of course, that the direct object NP is no longer adjacent to the verb that signals its functional role. As the direct object NP grows heavier, the tradeoff between cost and benefit changes so that heavy NP shift is more likely to occur. In cases where there is no direct object but there are instead two prepositional complements to a verb, as in (19), the heaviness constraint operates more freely, since both complements retain local signals of grammatical function, whatever their relative ordering.

- (19) a. I spoke to Bill about our serious concerns.
b. I spoke about Bill to our closet friends.

I noted earlier that in questions, the use of *do* is favored when simple inversion produces a piling up of clitics. I also noted that this effect is distinct from the effects based on adjacency of verb and object that are the primary concern in this paper. The effect is, however, plausibly related to processing considerations. The piling up of clitics produces a sequence of weak syllables. Since clitics are obligatorily unstressed, this sequence becomes awkward to pronounce and perhaps difficult to perceive accurately. Exactly how these factors work is not known; but it is known that configurational language with well-developed systems of clitics tends to restrict the number of them that are strung together.

I end this tentative exploration of psycholinguistic effects that may underlie the historical development under investigation with a brief remark about a potential processing effect that might be thought to play a role but that the evidence suggests is, in fact, not involved. The effect I have in mind is the potential ambiguity that subject-verb inversion introduces in questions, especially *wh*-questions. Consider the following sentence.

- (20) Which knight saw the King?

Before the rise of periphrastic *do*, (20) was structurally ambiguous between the two readings in (21):

- (21) a. (for which knight *x*) (*x* saw the king)
b. (for which knight *x*) (the king saw *x*)

This ambiguity is produced by the fact that Middle, like Modern English, did not show overt subject-verb inversion when the questioned constituent was the matrix subject. Therefore, the phrase *which knight* could either be the underlying subject of the verb, in which case *wh*-movement would not change the order of constituents and subject-verb inversion would not apply; or it could be the underlying object, in which case *wh*-movement would put the object in preverbal position, and subject-verb inversion would place the underlying subject in postverbal position. As the pair of sentences in (22) illustrates, the use of periphrastic *do* in these cases eliminates the ambiguity.¹³

- (22) a. Which knight saw the king?
b. Which Knight did the king see?

If this structural ambiguity played a role in promoting the *do* form, one would expect to see that role reflected in the differential distribution of *do* across linguistic environments. Thus, one would expect *do* to be more frequent in *wh*-question in direct object position than in other cases. The facts, however, show an opposite effect. Ellegård's data (1953:205) reveal that *wh*-questions are considerably less likely to include *do* than *yes/no* questions; furthermore, *do* is more likely to appear in adverbial *wh*-questions, as in (23), than in cases where the direct object is questioned, as in (24)

- (23) a. Why tel I this? (320 416/21)
b. Why do ye make this sorrow? (312 194/4)
(24) a. ...but what is our gest a doying, or what maketh our gest? (326 116/18)
b. But what like thing do you reade in all scripture of the single life? (338 42/38)

These facts are clearly impossible to square with the idea that avoidance of structural ambiguity plays an important role in the spread of periphrastic *do*.¹³

4. The competition among forms.

4.1 In the previous section I sketched the psycholinguistic assumptions necessary to account for syntactic drift on the basis of differential processing complexity among alternating forms. Of these assumptions, the basic one is that forms requiring more complex processing will be disfavored--all other things being equal--in comparison with less complex alternatives. It is fair to ask, however, why one should expect this.

The answer, I believe, can be seen if we model the communicative function of language as a stochastic process. We know that linguistic communication often fails for mechanical reasons. Noise, inattention, and other factors can cause listeners to fail to understand what their interlocutor is saying. When this happens, it may go unnoticed, but it is more likely to prompt listeners to ask for clarification or speaker to explain themselves after noticing spontaneously that something has gone amiss. Presumably, these communication failures are unevenly distributed across linguistic environments, with failure more likely to occur as load on the listeners' processing mechanism increases.

On this assumption one can construct a family of models of change. In a situation where only two forms are alternating, which are identical in meaning and discourse function, the models are simple and mathematically tractable. Suppose, for example, that the relative frequency of a pair of alternating, equivalent forms is fixed at any one time by community norm, i.e. by a tendency (which may or may not be specifically linguistic) for speakers to use a form at the same rate at which they hear it used. Some such assumption is needed to explain the fact, firmly established by quantitative studies, that speech communities share common rates of use for linguistic alternatives.

Suppose further that when a failure of communication occurs due to processing failure, the linguistic forms involved do not count as having been 'heard' by the listener. From these two assumptions it follows that if two alternatives differ in susceptibility to processing

failure, their relative frequencies as 'heard' by the speech community will be skewed toward the simpler variant by comparison with their relative frequencies as spoken. This skewing, so long as it is above a minimum threshold necessary for detection by the psychological mechanism that monitors frequency, is sufficient to guarantee that the favored form will eventually replace its alternative, no matter what their original relative frequencies.

Another model that will produce the same result is a 'learning' model. Under this account, failure in communication has the same effect as punishment for choosing the wrong alternative does in an avoidance learning experiment (Sternberg 1963). Speakers learn not to use the disfavored form because it leads more often to an undesired outcome--misunderstanding--than does use of the favored form. Because the frequency with which misunderstanding occurs is low, however, the changeover to the new form is very gradual. As with the first model, speakers born into the community adopt the frequency of use of the two forms characteristic of the community at the time when they acquire the language and modify it gradually as a result of their life experience.

In addition to the two models I have sketched, there are many other plausible ones that could be devised; we are certainly not in a position to choose among them with the data available. What is of interest is the fact that the models we are considering can all be expressed mathematically by the same well known function, the logistic, whose general form is given in (25).

$$(25) \quad p = \frac{1}{1 + ((1-p_0)/p_0)e^{st}}$$

When graphed against time, the logistic function has a characteristic S-shape, illustrated in Figure 2.

The logistic function expresses a number of basic growth relationships in population biology and genetics. Of most interest to us is the fact that it expresses the rate of replacement of one species for another in a context where the two compete with differential reproductive

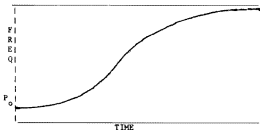


Figure 2. The logistic function.

success for the same resources (Crow and Kimura 1970), a situation that is exactly analogous to the replacement of one linguistic form by another when the two are unequally likely to reinforce their own future use.¹⁵

The value of assuming that the curve of replacement of one linguistic form by another is logistic is that one can then use sampling data to estimate the two parameters of the logistic equation, s and p_0 . Of these parameters, s represents the advantage of one form over the other and p_0 the initial relative frequency of the favored form. By comparing the values of these parameters across linguistic environments, one can obtain a quantitative description of the time course of a change, a description

whose theoretical interest I hope to demonstrate in the following discussion.

A word is in order at this point about the justification for assuming that the course of change will follow a logistic curve.¹⁶ While this equation follows from plausible assumptions, and while data of a linguistic change like the rise of *do* may appear upon inspection to follow a roughly S-shaped growth curve, we cannot demonstrate conclusively that the logistic rather than another function with a similar graph is the correct function to use. For one thing, the family of such functions is infinite and for another, the amount of data needed to choose among similar curves is prohibitive in a discipline dependent on observational data.

Fortunately, however, the very fact that these curves are similar means that choosing among them makes less difference than it might. The parameter estimates based on the assumption that the data follow the logistic curve are robust; i.e. they are likely to remain valid even if it is not the logistic function itself but some function with a similar shape that most accurately describes the empirical data in a particular case. This is especially true when the parameter estimates are used for comparison across subenvironments of a single change, for then any bias in the estimates is likely to be constant across the cases compared.

5. Explaining the rise of *do*.

5.1 One may now ask how the assumptions sketched in the previous sections can help us to understand the time course of the rise of *do* as represented in Ellegård's statistics and graphed in Figure 1. The graph reveals two features of special interest: the use of *do* rises in affirmative declarative sentences until 1560 but then falls gradually to zero in that environment, and the use of *do* in negatives falls for a short time after 1560 before it rises again and eventually becomes categorical. A specific account of this change should be able to explain these two characteristics of the time course of the change. It should also provide a test for my own attempt.

I shall begin by concentrating on the period up to 1560, during which there is a monotonic increase in the frequency of *do* in all environments, although some en-

vironments favor the new form much more than others. Under the assumption that the four curves are generated by a set of logistic functions, there are basically two models of the change that are possible. The first model, whose correctness we implicitly assumed in our own initial work (Kroch, Pintzuk and Myhill 1982) might be called 'the direct influence model.' It claims that when *do* first becomes available as a periphrastic auxiliary, it appears in all environments with an equal, very low frequency. Then, due to the shift to underlying SVO word order, it begins to rise in frequency at a different rate in each environment. It rises most steeply in negative questions because they have the most features that favor its use (see Kroch et al. for discussion), and, in general, the steepness of the rise reflects for each environment the strength of the processing advantage conferred by its use in that context.

The direct influence model entails that psycholinguistic pressures directly determine the rate of increase in the use of *do*. In particular, the fact that *do*, in all its environments, is introduced by a single grammatical rule does not lead change in one context to affect usage in other contexts. The grammar may change when the simple form disappears from some environments, but the form of its rules has no effect on the course of the change. The model entails this consequence because a grammatical analysis of periphrastic *do* in the period in question shows that it must have introduced by a single optional phrase structure rule (or some notational equivalent) and not by different rules in different environments.

Only this, the simplest grammatical analysis, can account in a straightforward way for several facts. (1) Periphrastic *do* occurs from the earliest period in affirmative declarative sentences as well as in those sentences where its use confers a processing advantage. (2) The use of *do*, until it becomes categorical, does not change the rules for question inversion and not placement, which need do no more than refer to the first (i.e., the tensed) verb in the clause. (3) Periphrastic *do* becomes assimilated to the syntactic class of the modals (Jackendoff 1977, Lapointe 1981), sharing with them the property of not appearing in the complement of any other verb.

One problem with the model I have described is that the steady increase in the use of *do* in affirmative declarative sentences up to 1560 becomes somewhat of a puzzle. Ellegård, who believes that the rise of *do* is motivated by functional effects (although he does not discuss that issue at length), recognizes the difficulty posed by the behavior of the affirmative declaratives. His solution is to say that the use of *do* in this environment, unlike in others, is a stylistic conceit. He considers it a literary fashion that gains adherents for a time and then becomes outmoded (165 ff.), accounting for its eventual disappearance from the language. Ellegård provides some—but not much—evidence for the correctness of his proposal; I have found no direct evidence to rule it out. We must conclude, therefore, that the 'direct influence model' is compatible with what is known about the behavior of affirmative declarative *do*, although it certainly sheds no added light on the question.

The second model for the change might be called the 'mediated influence model.' Under this model the influence of processing effects on changes in the frequency of syntactic alternants in different environments is constrained by the way that the forms are categorized by the grammar. If a form (here *do*) is introduced into a number of different environments by a single rule, then it must change its frequency of use in all environments. Psycholinguistic effects that cause one environment to favor *do* more than another will be reflected in a higher frequency of use of *do* in the favored environments, but change in frequency of occurrence affects all environments concurrently.

Such a model implies that the rate of increase in the use of *do* should be the same in all environments and that the degree to which each environment favors or disfavors the use of *do* should be the same at every point in time. In particular, it predicts that the use of *do* in affirmative declarative sentences should rise along with its use in other environments, even though there is apparently no psycholinguistic motivation for its use there. That fact, the lack of motivation is reflected in its much lower frequency in this environment; however, since the grammatical rule that introduces *do* in questions and negatives also permits its use in affirmative

declaratives, there is no way for the frequency of *do* to rise in the former contexts but not in the latter.

The mediated influence model implies that the speech community shares a norm for the overall rate of use of the *do* form, which is defined as single form by the fact that a single grammatical rule governs its distribution. This overall frequency of use is presumably adopted by speakers as they learn the language. It may be a part of grammatical knowledge, but it is only necessary to assume that it is active in sentence generation at whatever point choices among syntactic alternatives are made. The interesting property of the mediated influence model is not that entails that people have knowledge of frequencies¹⁷, but rather that what counts as a form whose frequency in usage can be controlled by community norms is defined by what the rules of grammar group as a single construction. Under this model, the distribution of *do* across environments is not learned. Rather, it is an automatic consequence of the operation of the processing constraints on behavior (see Kiparsky 1971 for a suggestion along these lines).

Just how these constraints would come to govern choices made in sentence generation is not known. Presumably, they would function indirectly by their influence on the monitoring that speakers do of their own output. The use of *do* would simplify this monitoring to different degrees in different environments, and speakers' use of the form would reflect these differences within the constraint on overall frequency imposed by the community. The model still allows the increase in use of *do* over time to be attributed to the processes described in section 4, but these processes would be constrained to influence only the overall rate of use of *do*, because this is the only frequency that language learners would be tracking and matching.

It is obvious that the two models I have presented are both based on many assumptions whose validity we have no way of demonstrating. Their utility, therefore, lies not in details of implementation but in the feature that distinguishes them; that is, whether the grammar's characterization of a form as the same (because introduced by a single rule) across a number of environments directly constrains how the use of the form can rise or fall. This question is clearly of central importance to the

theory of syntactic change, and it has crucial implications for synchronic linguistic theory as well. If the progress of syntactic changes is governed entirely by functional psycholinguistic and other factors external to grammar (e.g., style and social group membership), then studying the changing usage frequencies of the competing forms cannot be expected to help reveal the organization of grammatical knowledge.

Moreover, if such a result can be maintained in the explanation of diachronic processes, the plausibility of associating usage frequencies to grammatical constructs in synchronic analyses is greatly reduced. Under these circumstances, the divorce between the grammar and the performance system that governs choice among grammatical possibilities becomes complete. Each can be studied independently of the other, and the results obtained from the investigation of one will have little bearing on our understanding of the other.

On the other hand, if the progress of syntactic change as reflected in changing usage patterns turns out to be constrained by the grammatical organization of these forms, more interesting possibilities present themselves. In particular, it becomes possible to infer from the fact that a group of related syntactic constructions, because they change their frequencies of occurrence in tandem, are generated by a single rule. Moreover, if one of these constructions begins at a certain point to change independently from the others, it follows that the grammar must have changed. Thus, a change in patterns of usage becomes evidence for grammatical reanalysis. In the case of *do*, being able to use such reasoning would be especially attractive, as it might lead to an explanation of the change in behavior of affirmative declaratives and negatives that occurs after 1560.

Given the significance of the difference between a direct influence and a mediated influence model of change, I would like very much to find evidence that chooses between them. I believe I have found such an indication, favoring, for the case of periphrastic *do*, the mediated influence model; I will present the evidence in the paragraphs to follow. To the extent that this result can be confirmed and extended by further research, it promises, on the grounds I have given here, to provide a new paradigm for the relationship between diachronic

syntax and the theory of grammar, and also to provide a field for the investigation of processing constraints through the statistical study of usage data. While it is certainly early days to accept the validity of these proposals, I hope that the evidence to be presented will be sufficiently convincing to stimulate further work along the lines I have sketched.

The reason why it is possible to find evidence in the data on *do* that distinguishes between our two models of syntactic change is that the models make different predictions about the relationship between the curves of increase in the various environments. Under the direct influence model one would expect the rate of increase of the curves to be different for each curve, with the rates reflecting the degree to which each environment favors the *do* form. This prediction is an obvious one that follows from the fact that in the direct influence model, the strength of the processing effects directly determines the rate at which the *do* form replaces its alternant in each environment. Under the mediated influence model, on the other hand, one expects the *s* parameter to be the same for all the curves, since the rate of increase in this model is the same for all environments.

Without a mathematical model of the curves, these predictions would, of course, be difficult to evaluate, as it would not be clear how to measure their rates of increase. Visual inspection of Figure 1 suggests that the rate of increase is different for each environment. This appearance is, however, deceptive because it depends on assuming that the average slope of the curves is a reasonable measure of the rate of increase, i.e., that the curves are essentially linear. Both examination of the curves and experience with curves describing replacement of one form by another suggest, however, that this assumption is false. If we assume instead that the curves are logistic, we can calculate a more reasonable estimate of their rate of increase.

The parameter that determines how fast a logistic curve increases with time is *s* parameter in equation (25). When percentage data that follow a logistic equation are converted into their logistic transforms (where the logistic $(p) = \ln(p/(1-p))$), the transformed data lie

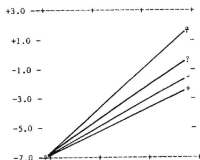


Figure 3. Predicted graph of the transformed frequencies of the rise of *do* under the direct influence model.

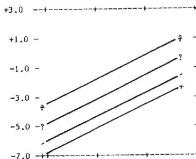


Figure 4. Predicted graph of the transformed frequencies of the rise of *do* under the mediated influence model.

on a straight line with slope equal to the parameter s and a y -intercept of p_0 . Thus, if we transform the percentage data in Table 1 and obtain an estimate of the parameters of each curve by fitting the transformed data to a regression line, we can choose between the two models under discussion.¹⁸ The direct influence model predicts that the regression lines should all have different slopes and intercept the y -axis at a common point early in the historical development. This point corresponds, of course, to the point in time when the processing effects responsible for the rise of *do* begin to take effect. Figure 3 represents this outcome.

The mediated influence model, on the other hand, predicts that the slopes of all the lines should be equal and hence that the lines should be equidistant at every point in time. Such a relationship corresponds to the hypothesis that the processing effects on the frequency of *do* in different environments are constant across time.¹⁹ This outcome is represented in Figure 4.

Figure 5 shows the actual regression lines for the four environments calculated from Ellegård's data.

It is clear from inspection that the lines are essentially parallel. Moreover, a test of the statistical significance of the slight differences in slope among the lines shows that the probability that the difference of this size is due to chance is greater than .30.²⁰ I conclude, therefore, that the direct influence model may be rejected with some confidence and that the mediated influence model has achieved a substantial measure of confirmation.

5.2 Having provided evidence that the increase in frequency of the use of *do* up to 1560 in all environments reflects an increase in the application of a single grammatical rule that introduces *do*, I will now briefly explore the consequences of this result for purposes of understanding the course of the change after 1560.

Examination of Figure 1 shows that after 1560 the curves no longer move in tandem. While *do* in affirmative questions continues to increase in frequency along roughly the same path as before, in affirmative declaratives it begins a monotonic decline toward zero. The behavior of *do* in negatives, both declaratives and questions, is more complex, as the curves for these environments

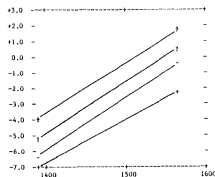


Figure 5. Actual regression lines fitted to the transformed frequency data for the rise of *do*.

decline in tandem before rising toward 100 percent. While I cannot claim to be able to account for this pattern in all its details, its main features are amenable to at least a tentative explanation if one assumes the validity of the mediated influence model of the change.

The decline in affirmative declarative *do* presumably reflects the fact that processing factors do not favor its use. But this decline is only possible under the mediated influence model if there is a grammatical reorganization at the point of inflection for the curve. If affirmative declarative *do* declines toward zero while *do* in affirmative questions continues to increase in frequency, it must be because different rules now govern the two environments.

There are several changes in the grammar that would free *do* in affirmative declaratives to decline independently of the other environments, and they have different consequences for the rest of the system. The simplest change would be the addition of an optional rule that deleted unstressed *do* when it appeared immediately before another verb in the surface string.²¹ The decline in the

use of *do* in affirmative declaratives would then reflect an increasing frequency application of this rule, which would have no effect on other environments. If the curves for the negative environments did not decline for a time along with the curve for affirmative declaratives, this formulation of the reanalysis might be adequate; but under the actual circumstances, I would prefer a formulation that accounts for the behavior of the negatives as well as the other curves. Such a formulation is, of necessity, more complex than the mere addition of a *do* deletion rule.

Let us suppose that the grammatical reanalysis that occurs in the middle of sixteenth century involves the introduction of a new rule of subject-auxiliary inversion in questions, with *do* support of any consequent stranded tense, instead of a *do* deletion rule. The introduction of a subject-aux inversion rule would change the way the grammar grouped forms, since the *do* in questions would now have a different grammatical source from the *do* in declaratives, both affirmative and negative. Use of *do* in questions would continue to rise as the output of subject-aux inversion completed against the output of the old subject-verb inversion rule, which would eventually disappear.

Since the processing factors favoring *do* in negative sentences are weak (due to the clitic status of *not*), the removal by the original phrase structure rule of questions from the set of environments receiving *do* might change the balance of processing factors enough for the use of this *do* to be disfavored overall, and so account for its decline. *do* would decline both in negative sentences and affirmatives since the grammar would continue to categorize them together. At a later point, a rule placing *not* after the first auxiliary verb in a sentence (again with *do* support for stranded tense) would be introduced. Once this rule was available, the frequency of *do* in negative sentences would be free to rise again, and affirmative declarative *do* could continue its decline alone.

The unsolved problem in this story of reanalysis is, of course, why the reanalyses occur when they do, rather than earlier or later. We have no satisfactory solution to this problem; our proposals will remain unconfirmed until and unless further empirical investigation sheds

light on the issue. There are, however, plausible avenues for further investigation that give us hope of eventually making progress. The two questions that need answers are, Why does reanalysis occur in 1560 and why does *do* in negative sentences rise again in frequency after falling with affirmative declaratives for a generation? As far as the first question is concerned, our best guess is that the introduction of the subject-aux inversion rule is triggered when the type of question that most favors *do*, the yes/no question with full NP subject, reaches essentially 100 percent use of the *do* form. Under this circumstance, language learners project a grammar that makes the *do* form obligatory in this environment.²²

The new grammar must introduce subject-aux inversion and *do* support (or some other grammatical innovation with equivalent effect) for this subset of questions because subject-verb inversion is no longer possible. The new grammatical analysis is then taken to be the source of *do* in all questions, except perhaps negative questions, whose curve patterns with that of the negative declaratives. In this way, the question environment is separated from the other environments, leaving them to change independently. Concerning the second question, on the behavior of negatives, I have only one brief remark to make. The fact that negative questions pattern with negative declaratives in the period after 1560 indicates that negative questions could not have continued to be formed with *do* if it had dropped out of use in negative declaratives. For reasons we have yet to understand, the reformulation of *not* placement seems to have been triggered in order to preserve the syntactic parallel between affirmative and negative questions.

6. Conclusion.

What I have attempted in this discussion is certainly an overambitious synthesis of our still woefully imperfect knowledge of three areas that are not often brought together in contemporary linguistics: the study of syntactic change, the theory of linguistic variation, and the psychological theory of human sentence processing. I have no doubt that many of the assumptions I have made in order to carry the discussion forward will have to be modified or abandoned as our knowledge in each

of these areas increases. Nevertheless, I believe that I have succeeded in sketching a framework that allows us to use quantitative data on usage to illuminate historical change. Most important, this framework allows us to analyze syntactic drift in detail and to account for it without denying its gradual or incremental character. I have also managed to obtain at least one substantial empirical result that bears on the proper statement of the relationship between patterns of usage and rules of grammar: during the period of monotonic increase, the spread of periphrastic *do* is a unitary phenomenon. Its rate of use rises across all environments in such a way that the relationships among environments remain fixed as the overall rate of use increases.

From this pattern and from the fact that a single grammatical rule can account for the appearance of periphrastic *do* in all its distributional contexts, I am able to propose that the categorization of a set of forms by the grammar as instances of a single construction constrains them to change in tandem. If this hypothesis can be corroborated by other studies, then I will have achieved a goal worth striving for: the integration of quantitative and discrete (grammatical) analysis within a single theory of language. Even if I have left my readers skeptical as to the plausibility of my specific proposals, I hope that I have convinced them that further work along the lines I have been following should be on the agenda for future research.

NOTES

My warm thanks go to all those with whom I have worked on the problem of *do* over the past year and more. I am especially grateful to David Sankoff of the Centre de Recherches en Mathématique Appliquée of the Université de Montréal, who arranged a visiting research appointment for me at the Center in the summer of 1982 and worked with me on the problem of modeling the time course of the change. The basic result reported in this paper was worked out in discussions between the two of us. I also owe many thanks to Susan Pintzuk of the Linguistics Department at the University of Pennsylvania, who has been working with me on the linguistic problems raised by the history of *do*, and to John Myhill, also of our depart-

ment, who helped on the project in the spring of 1982. Finally, I thank William Labov for several extremely helpful discussions and the members of my fall 1982 research seminar for serving as a critical sounding board for an earlier version of the analysis presented here.

1. The examples in this paper are all taken from sources used by Ellegård, although not all the specific citations were considered by him in his analysis. For convenience, we have indicated the source of example sentences in parentheses, giving first the number assigned to the source by Ellegård, followed by the page and line numbers of the quotation.

2. The reanalysis we are carrying out is made much easier by two factors. First, Ellegård's quantitative analysis is thorough and insightful, thus providing a solid basis for further work. Second, he has taken the extraordinary pains to publish the page and line numbers for all the tokens he included in his analysis, allowing us to go back to the original texts without having to collect the data over again from scratch. Ellegård's thoroughness and scholarship establish a standard for research that those of us who work in quantitative linguistics would do well to emulate.

3. As is the case with most investigations in historical linguistics, the material of Ellegård's study is limited to the literary language, although the conclusions he draws are meant to apply to the development of the language as a whole. This mismatch is as unavoidable as it is dangerous. In the case at hand, however, we have reason to think that the literary and spoken language may not differ significantly, because there is little evidence of stylistic conditioning in the use of periphrastic *do*.

4. If the slow drift of main verb *have* into the class of verbs that take *do* is considered part of the change, it might be said that in many dialects the change has yet to go to completion.

5. Hausmann attempts to account for the rise of periphrastic *do* as a sequence of rule reorderings and changes in the conditions on various transformations, basing his analysis on Aspects-style grammar of the auxiliary. From our perspective his account is entirely unsatisfactory as it offers no explanation for why the change is so gradual

and none for its stratification by linguistic environment.

6. As the reader will see, the processing simplification induced by the use of *do* is in the parsing rather than the generation of sentences; yet the data from which we infer the existence of this parsing effect are necessarily so in a historical study-data of production. To give a coherent account of this data, we clearly must postulate the existence of a feedback relationship between generation and parsing so that the difficulties of the latter process come to guide the operation of the former. That this feedback should exist is to be expected, however, since we know from the data of self-correction that speakers monitor the grammatical structure of their productions.

7. These data demonstrate Ellegård's point only for the eastern dialects, in which causative *do* is widely used until the fifteenth century. The data from the western dialects, in which the periphrasis first appears, is harder to interpret because the periphrastic instances already occur in the earliest texts and causative *do* is never common. Ellegård uses a number of complex indirect arguments to support his claim that the development in the west must have been comparable to that in the east; but it is also possible that periphrastic *do* came into being because *do* was borrowed by the western dialects from the east. Since the western dialects contained the productive causative verbs *make* and *let*, the borrowed form underwent a semantic mutation in the process of being adopted. The initial appearance of periphrastic *do* in the east could then be attributed to that dialect's borrowing the form back with the changed meaning after causative *do* has receded in favor of *make*.

8. We should note here that the *do* in affirmative declarative sentences that Ellegård has charted is unemphatic and unstressed. He excludes those cases in which context suggests that *do* is being used to convey emphasis or insistence, since that *do* in that environment eventually becomes obligatory. Unstressed *do* is completely absent from contemporary English usage, although memory of it is preserved in certain archaic legal formulae like 'I do hereby swear....'

9. This connection, however, is not established by Ellegård's work, and further statistical analysis would be

needed here to determine how adverb placement is related to other features of word order.

10. Historically, *not* is an emphatic negative particle that co-occurred with the preverbal clitic *ne*, analogous to the French postverbal negative *pas*. In the course of Middle English, the use of *ne* declines and spelling evidence indicates that *not* loses its emphatic character and becomes an unstressed particle. By the time that periphrastic *do* becomes important, the use of *ne* is rare.

11. The word order factors that influence the use of *do* in questions are summarized in the following table, modified from Kroch *et al.*

The effect of transitivity on the likelihood of *do* in questions 1490-1600

SUBJECT	OBJECT	% DO	PROB <i>do</i>	N
noun	pro	95	.96	40
noun	noun	93	.87	68
noun	intrans	55	.45	87
pro	pro	65	.62	168
pro	noun	40	.28	268
pro	intrans	35	.23	433

12. Needless to say, parenthetical phrases can occur between verb and object in modern English, so that a sentence like (i) is perfectly acceptable:

(i) John bought, I heard, several books yesterday.

As Mark Baltin has pointed out to me, however, parentheticals can appear anywhere in a sentence, even between the determiner and the head of an NP, as in (ii):

(ii) The, I think, biggest tragedy was the lost opportunity.

These facts show that the placement of parenthetical phrases is not a matter of sentence grammar at all and must be accounted for by an entirely different component of the language system.

13. A similar, though less general, ambiguity arises in certain cases, where yes/no questions without *do* become identical as strings to imperatives, as in the following case:

(i) Ask the men to leave?/!

The parsing problem here is similar to one discussed by Marcus (1980:207 ff.), who points out that sentences like (ii) are potentially ambiguous:

- (ii) Have the eggs fried?/!

These examples, of course, may not pose a problem since intonation would disambiguate them under normal circumstances. Nonetheless, it is unclear whether the use of intonational information to establish constituent structure in such cases should be considered not to add to processing load.

14. Marcus (personal communication) points out that there is another problem in parsing that might play a role in the change--namely, that of assigning lexical category membership to the words of a sentence as they are encountered. Since so many English words can be used as either noun or verb, and since English morphology is so impoverished, lexical category assignment is indeed a difficult problem for parsing. Unfortunately, it is not obvious how the use of *do* would simplify that problem in the sentences in which it appears.

15. Several historical linguists, among them Weinreich, Labov and Herzog (1968:113) and Bailey (1973:77), have remarked that the time course of linguistic change seems to follow an S-shaped curve. Tony Naro, moreover, recently brought to my attention an article by Altmann, von Buttlar, Rott, and Strauss (1983), in which the idea of modeling this curve with the logistic function is discussed.

16. The remarks in this paragraph come from discussions between David Sankoff and myself of the statistical issues involved. Responsibility for the formulation I have given them is, however, entirely mine.

17. As Labov has pointed out to me, it may also be necessary to postulate knowledge of frequencies in the direct influence model to account for why the use of the *do* form does not immediately become universal once processing factors begin to favor it. Under that model, speakers will eventually track the frequency of *do* separately in each environment. If the model were confirmed empirically, this would raise the issue of how environments were defined, since one would not want to say that every distinguishable context constituted a separate environment for which usage frequencies were independently tracked.

18. To obtain this fit, I used an algorithm written for me by David Sankoff. The procedure used is much like simple linear regression except that an iterative maximum likelihood calculation is used to define 'best fit' instead of the standard least squares approach.

19. It is interesting to note that variable rule analysis also makes use of the logistic transform of percentage data to calculate factor effects. Because of this, a situation in which the regression lines for the different environments are parallel corresponds to one in which the factor effects of a series of variable rule analyses carried out at different points in time are constant, with only the input probability for the rule increasing with time.

20. The test of significance used to determine this Figure is the same as the one used in variable rule analyses. It is based on the difference on log likelihood between two runs of the curve fitting program, one of which allows the slopes of the curves to vary independently and the other of which forces them to be identical. The difference in log likelihood, when multiplied by 2, behaves like a chi-square statistic, with degrees of freedom equal to the difference in the number of parameters free to vary independently under the two runs. In this case the difference in log likelihood was 1.75 and there were 3 degrees of freedom.

21. Needless to say, this deletion rule would have to apply after subject-verb inversion and *not* placement. In a grammar without rule ordering of the type assumed in Chomsky and Lasnik 1977 and later work in that tradition, the rule is guaranteed to apply at the right point in the derivation because, like other deletion rules, it belongs to the phonological component, which applies to the output of the transformational component.

22. At this stage, the dying English rule of subject-verb inversion looks like the modern French rule, which allows inversion of subject and main verb in all environments except yes/no questions with full NP subjects. The parallel is striking; the fact that French is, like English, a strict SVO language suggests that a quantitative investigation of the competition in questions between the use of inversion and the question particle *est-ce que* should prove interesting.

REFERENCES

- Altman, G., H. von Buttlar, W. Rott, and W. Strauss. 1983. A law of change in language. In: B. Brainard, (ed.), *Historical Linguistics. Quantitative Linguistics*, Vol. 8. Bochum, West Germany: Studienverlag Dr. N. Brockmeyer.
- Bailey, C-J. N. 1973. *Variation and Linguistics Theory*. Arlington, VA: Center for Applied Linguistics.
- Bock, J. K. 1982. Toward a cognitive psychology of syntax. *Psychological Review*. 89:1-47.
- Canale, M. 1978. Word order Change in OE: Base Re-analysis in Generative Grammar. Ph.D. dissertation, McGill University.
- Carroll, J. 1981. On fallen horses racing past barns. In: Masek et al. (eds.), *Papers From the Parasession on Language and Behavior*. Chicago: Chicago Linguistics Society.
- Chomsky, N., and H. Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8:425-504.
- Crow, J., and M. Kimura. 1970. *An Introduction to Population Genetics Theory*. New York: Harper and Row.
- Ellegård, A. 1953. *The Auxiliary do: The Establishment and Regulation of its Use in English*. Stockholm: Almqvist & Wiksell.
- Engblom, W. 1938. *On the Origin and Early Development of the Auxiliary Do*. Lund: C. W. K. Gleerup.
- Frazier, L, and J. D. Fodor. 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6:291-325.
- Fries, C. C. 1969. On the development of the structural use of word order. In: Lass (ed.), *Approaches to English Historical Linguistics*, Holt, New York: Rinehart and Winston.

- Haiman, J. 1974. *Targets and Syntactic Change*. The Hague: Mouton.
- Hausmann, R. B. 1974. The development of periphrastic do in English. In: J. M. Anderson and C. Jones (eds.) *Historical Linguistics*, Vol. 1. New York: American Elsevier.
- Jackendoff, R. 1977. *X Syntax: A Study of Phrase Structure*. Linguistic Inquiry Monograph No. 2. Cambridge: MIT Press.
- Jacobson, B. 1951. *Inversion in English*. Uppsala: Almqvist and Wiksell.
- Kiparsky, P. 1971. Historical linguistics. In: W. O. Dingwall (ed.), *A Survey of Linguistic Science*. Linguistics Program University of Maryland.
- Kroch, A., S. Pintzuk, and J. Myhill. 1982. Understanding do. In: K. Tuitt et al. (eds.), *Papers from the 18th Regional Meeting*. Chicago: Chicago Linguistics Society.
- Lakoff, R. 1972. Another look at drift. In: R. Stockwell and R. K. S. Macaulay (eds.), *Linguistic Change and Generative Theory*. Bloomington: Indiana University Press.
- Lapointe, S. 1981. A lexical analysis of the English auxiliary verb system. In: T. Hoekstra et al. (eds.), *Lexical Grammar*. Dordrecht: Foris Publications.
- Lightfoot, D. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
- Malkiel, Y. 1981. Drift, slope and slant. *Language* 57:535-70.
- Marcus, M. 1980. *A Theory of Syntactic Recognition for Natural Language*. Cambridge: MIT Press.

- Mossé, R. 1952. *A Handbook of Middle English*. Baltimore: Johns Hopkins Press.
- Rousseau, P., and D. Sankoff. 1978. Advances in variable rule methodology. In: D. Sankoff (ed.), *Linguistic Variation: Models and Methods*. New York: Academic Press.
- Rydén, M. 1979. *An Introduction to the Historical Study of English Syntax*. Stockholm: Almqvist & Wiksell International.
- Sapir, E. 1921. *Language: An Introduction to the Study of Speech*. New York: Harcourt, Brace.
- Schmidt, D. 1980. *A History of Inversion in English*. Ph.D. dissertation, Ohio State University.
- Sternberg, S. 1963. Stochastic learning theory. In: R. D. Luce et al (eds.), *Handbook of Mathematical Psychology*, Vol. II. New York: Wiley and Sons.
- Sweet, H. 1899. *A new English Grammar*. Oxford: Clarendon Press.
- Wanner, E. 1980. The ATN and the sausage machine: Which one is baloney? *Cognition* 8:209-25.
- Weinreich, U., W. Labov, and M. Herzog. 1968. Empirical foundations for a theory of language change. In: W. Lehmann and Y. Malkiel (eds.), *Directions for Historical Linguistics*. Austin: University of Texas Press.

Since 1986 the following titles have been published or are scheduled for publication in 1989:

43. AKAMATSU, Tsutomu: *The Theory of Neutralization and the Archiphoneme in Functional Phonology*. Amsterdam, 1988. n.y.p.
44. JUNGRAITHMAYR, Hermann & Walter W. MUELLER (eds): *PROCEEDINGS OF THE FOURTH INTERNATIONAL HAMITO-SEMITIC CONGRESS*. Amsterdam, 1987.
45. KOOPMAN, W.F., F.C. VAN DER LEEK, O. FISCHER & R. EATON (eds): *EXPLANATION AND LINGUISTIC CHANGE*. Amsterdam, 1987.
46. PRIDEAUX, Gary D., and William J. BAKER: *STRATEGIES AND STRUCTURES: The Processing of Relative Clauses*. Amsterdam, 1986.
47. LEHMANN, Winfred P.: *LANGUAGE TYPOLOGY 1985. Papers from the Linguistic Typology Symposium, Moscow, 9-13 Dec. 1985*. Amsterdam, 1986.
48. RAMAT, Anna Giacalone (ed.): *PROCEEDINGS OF THE VII INTERNATIONAL CONFERENCE ON HISTORICAL LINGUISTICS, Pavia 9-13 September 1985*. Amsterdam, 1987.
49. WAUGH, Linda R. & Stephen RUDY (eds): *NEW VISTAS IN GRAMMAR: Invariance and Variation*. Amsterdam/Philadelphia, 1989. n.y.p.
50. RUDZKA-OSTYN, Brygida (ed.): *TOPICS IN COGNITIVE LINGUISTICS*. Amsterdam/Philadelphia, 1988.
51. CHATTERJEE, Ranjit: *Aspect and Meaning in Slavic and Indic*. Amsterdam/Philadelphia, 1989.
52. FASOLD, Ralph & Deborah SCHIFFRIN (eds): *LANGUAGE CHANGE AND VARIATION*. Amsterdam/Philadelphia, 1989.
53. SANKOFF, David (ed.): *DIVERSITY AND DIACHRONY*. Amsterdam, 1986.
54. WEIDERT, Alfons: *Tibeto-Burman Tonology. A Comparative Analysis*. Amsterdam, 1987.
55. HALL, Robert A. Jr.: *Linguistics and Pseudo-Linguistics*. Amsterdam, 1987.
56. HOCKETT, Charles F.: *Refrubishing our Foundations. Elementary Linguistics from an Advanced Point of View*. Amsterdam, 1987.
57. BUBENIK, Vít: *Hellenistic and Roman Greece as a Sociolinguistic Area*. Amsterdam/Philadelphia, 1989.
58. ARBEITMAN, Yoel L.: *FUCUS. A Semitic/Afrasian Gathering in Remembrance of Albert Ehrman*. Amsterdam/Philadelphia, 1988.
59. VOORST, Jan van: *Event Structure*. Amsterdam/Philadelphia, 1988.
60. KIRSCHNER, Carl and Janet DECASARIS (eds): *STUDIES IN ROMANCE LINGUISTICS*. Amsterdam/Philadelphia, 1989. n.y.p.
61. CORRIGAN, Roberta, Fred ECKMAN and Michael NOONAN (eds): *LINGUISTIC CATEGORIZATION*. Amsterdam/Philadelphia, 1989. n.y.p.
62. FRAJZYNGIER, Zygmunt (ed.): *CURRENT PROGRESS IN CHADIC LINGUISTICS*. Amsterdam/Philadelphia, 1989. n.y.p.

A full list of titles published in this series is available from the publisher.