

North East Linguistics Society

Volume 25 *Proceedings of the North East Linguistic Society 25 -- Volume Two: Papers from the Workshops on Language Acquisition & Language Change*

Article 11

1995

Notions of Economy in Language Change: The Spread of Periphrastic 'Do'

Mark D. Arnold
University of Maryland

Follow this and additional works at: <https://scholarworks.umass.edu/nels>



Part of the [Linguistics Commons](#)

Recommended Citation

Arnold, Mark D. (1995) "Notions of Economy in Language Change: The Spread of Periphrastic 'Do'," *North East Linguistics Society*. Vol. 25 , Article 11.

Available at: <https://scholarworks.umass.edu/nels/vol25/iss2/11>

This Article is brought to you for free and open access by the Graduate Linguistics Students Association (GLSA) at ScholarWorks@UMass Amherst. It has been accepted for inclusion in North East Linguistics Society by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

Notions of Economy in Language Change: The Spread of Periphrastic ‘Do’

Mark D. Arnold

University of Maryland

Within Chomsky’s (1993) Minimalist Program for Linguistic Theory (MPLT), the role of simplicity in the theory of generative grammar takes on new significance: rather than looking to a simplicity metric as a tie breaker for empirically equal proposals, the notion of simplicity becomes a theoretical primitive. In other words, Principles of Economy are taken to define the range of well-formed linguistic structures rather than our preference for one analysis over another. The formalization of the role of simplicity in the theory inevitably provides a certain naturalness to proposals which invoke simplicity in the analysis of some linguistic paradigm, especially in the realm of language change, where the assumption that languages change to become “simpler” has been a traditional view for centuries. The aim of this paper is to reconsider the role of simplicity in the process of language change.

As a case study, I discuss the spread of periphrastic ‘do’ in English. In addition to considering two ways in which (some version of) simplicity falls short in explaining the spread of ‘do’, I also provide an account of the nature of the relationship between the use of ‘do’ and the loss of verb movement in English. In particular, I argue that a change in the abstract Case system was a central force in the spread of ‘do’ and the loss of verb movement. Empirical motivation for the proposal comes, in part, from certain facts concerning the early use of ‘do’, and more generally from the strong chronological parallels between the spread of ‘do’, the innovation of indirect object passives, the development of ‘that-trace’ effects, and the loss of verb movement. Furthermore, in looking at the problem from a minimalist perspective, I propose that viewing ‘do’ as a “first resort” allows us to use the historical pattern of ‘do’ in affirmative declaratives as a means of dating the loss of verb movement as a productive mechanism.

In section 1, I briefly outline the role of economy in the discussion of Roberts (Spring 1994 lectures, University of Maryland, based on Clark and Roberts (1993)), and Kroch (1989). The short-comings of both approaches will be clear given the range of facts discussed in section 2. Section 3 examines the idea that ‘do’ should be seen as a “first resort”. The final section has two parts, the first of which is a brief digression motivating a modification to the minimalist framework. In turn, the modification allows for an analysis of the spread of ‘do’ to follow from economy considerations.

1. Extra-Grammatical Simplicity in Language Change

1.1 Acquisition and Change

Clark and Roberts’ (1993) Genetic Algorithm (GA) allows for a very explicit characterization of diachronic change, including the change under study here, as discussed by Roberts (Spring 1994 lectures). With respect to the history of English, the GA makes precise the way in which sentences with modals or ‘do’ could be produced by a speaker with a verb movement grammar, but could be re-analyzed by the language learners as being instances of non-verb-movement. The summary below outlines Roberts’ proposal that two preliminary changes in the input set the stage for the GA’s fitness metric to hypothesize a non-verb-movement grammar.

First, the loss of a fully inflected agreement paradigm allowed the acquisition device to entertain the possibility that the V-features in I could be weak. This first step hinges on the assumption that rich morphological paradigms entail strong features (in MPLT terms), and that a less-than-fully-inflected paradigm can entail weak features. In addition to the weak Inflection features, the fact that the modals had become a grammaticalized category allowed structures with modals or ‘do’ to be “P-ambiguous”, which is to say that two different grammars were equally accurate at parsing strings containing modals or ‘do’. For example, *John did not leave* could be parsed by the grammar invoking verb movement by assuming that ‘do’ was base generated in VP, (as was arguably the case for the older generation,) or it could be parsed by the grammar which analyzed ‘do’ as a lexicalized inflectional element, and thus not be a trigger for a grammar with verb movement.

Recall the strategy of the acquisition device within the GA model: given two grammars which perform equally well with respect to parsing, the device invokes a simplicity metric which more highly values the grammar which requires the fewest nodes in its parse of some input. For the case in question, the grammar that doesn’t invoke verb movement would be more highly valued than the grammar with verb movement because postulating verb movement would require more nodes in the parse of the input, and so under this view, verb movement was lost due to the fact that the acquisition device is built to prefer a non-movement hypothesis if it is possible.

However, even though the GA approach can provide a precise formalization of how certain structures could come to be reanalyzed, it actually doesn’t provide any insight about why verb movement was lost for the following reason. The node-counting sub-routine of the fitness metric (which ultimately favors the non-movement grammar) is significantly less central to the system than the basic concern for providing a parse with fewest violations of UG. Thus, as long as children were hearing sentences with clear instances of overt verb movement, then the GA would continue to converge on a verb movement grammar because the non-movement grammar would fail to parse the inputs which were instances of overt movement.

The GA preference for a non-movement grammar would only come into play at a point when the number of P-ambiguous structures significantly outnumbered the structures

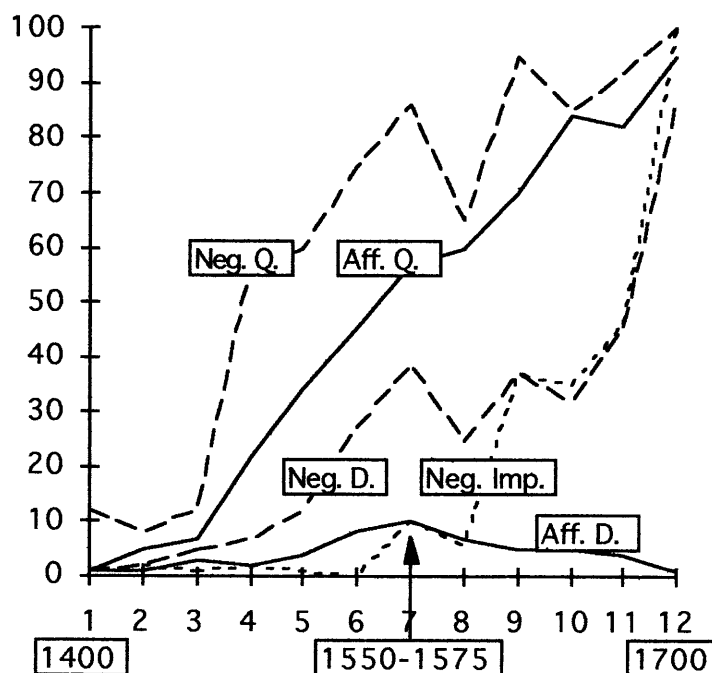
with overt verb movement, and since the model has nothing to say about why the input would have changed in such a way that the ambiguous structures would come to outnumber the unambiguous structures, it becomes clear that building a notion of simplicity into the acquisition device doesn't offer explanatory insight about the loss of verb movement in English.

1.2 Simplicity in Processing

Kroch (1989) provides an account for why the number of ambiguous structures would have increased significantly, namely, the use of 'do' alleviated a processing burden which had developed due to the collapse of the morphological case paradigm. In other words, once the overt case marking of NPs had been lost, then moving the verb out of the VP (potentially) created string discontinuity between the predicate and its complement(s); however, given the option of using 'do', the verb could remain in the VP, thus avoiding an unnecessary increased burden on the processing system.

The motivation for Kroch's proposal comes in large part from the findings of Ellegard (1953) regarding the use of 'do' in different sentence types, as follows. First, Ellegard discovered that during the period when the use of 'do' was spreading through the language, 'do' was more likely to be used with transitive verbs than with intransitives. Furthermore, Ellegard found that 'do' was statistically most likely to occur in negative questions, then in affirmative questions, then in negative declaratives, and was least likely to occur in affirmative declaratives, even in the period when it was presumably grammatical in unemphatic contexts. The chart in Figure 1, based on Ellegard's Table 7, represents the progress of 'do' in these sentence types:

Figure 1: Percent 'do' in different sentence types



There are two details worth noting at this point, given that they will play a role later in the discussion. First, note the down-turn in the use of 'do' in affirmative declaratives after 1575; I will argue that this detail can be taken to be an indication of the loss the productive use of verb movement. Second, consider the fact that the use of 'do' in negative

imperatives is less than the use of 'do' in affirmative declaratives until period 9, at which point 'do' in negative imperatives aligns essentially identically with the degree of frequency in negative declaratives; I take this pattern to be an indication that the change in the status of the negative marker from adverb to inflectional head was not complete until late in the 16th century.

More generally, the patterns found by Ellegard make sense given Kroch's proposal: if there is no object, as with intransitive verbs, then there is no need for the verb to stay in the VP in order to stay adjacent to its complement. Likewise, moving a verb out of the VP in a negative question creates more string discontinuity between the verb and a complement than moving a verb out of the VP in, e.g. a negative declarative; therefore, the processing burden would be highest in negative questions, and thus 'do' was most likely to be used in negative questions.

Kroch found further support for the processing account by finding that when subjects were pronominal, the use of 'do' in questions was less likely than when subjects were full NPs, thus supporting the idea that when more material (would have) intervened between a moved verb and its complement, 'do' was more likely to be used rather than overt verb movement. However, while the processing account allows us to make sense of certain aspects of the use of 'do', there are other developments about which the simplicity-in-processing account of 'do' provides no insight, and which suggest that a grammatical (rather than an extra-grammatical) explanation is in order.

2. Grammatical Correlations

In this section I provide the empirical motivation for looking at the spread of 'do' and the loss of verb movement as the consequence of a grammatical factor involving abstract Case. In short, the exceedingly close chronological parallels between the spread of 'do', the innovation of indirect object passives, the development of obligatory 'that-trace' effects, and the loss of verb movement suggest that we look for an explanation based on a single grammatical development.

Furthermore, while I do not adopt Kroch's processing account for the spread of 'do', I follow his intuition that Ellegard's patterns can be explained in terms of considering the relationship between a (potentially) moved verb and its complement(s); however, rather than assuming that 'do' alleviated a processing burden, I propose that a change in the Case system led to situation where the use of 'do' allowed for more economic derivations, a proposal I develop in sections 3 and 4. Sections 2.1 and 2.2 provide the empirical motivation for the analysis.

2.1 Indirect Object Passives

The innovation of structures in which an indirect object becomes the nominative subject of a passive, e.g. *He was given a book*, occurred in Middle English. Denison (1993:110), in summarizing the progress of indirect object passives (or "indirect passives") notes that the "first reasonably clear examples . . . are late 14th century" and that "they remained rare until late in the 15th century". Given these dates for the innovation and spread of indirect passives, Ellegard's (1953:209) summary of the innovation and spread of 'do' is striking: "prose examples begin to appear around 1400" and it is "not until the end of the 15th century that the do-form becomes widely used in prose texts".

So both 'do' and indirect passives begin to appear at the end of the 14th century, but neither innovation is widely used until the end of the 15th century. Thus it is not only the case that the innovations of the forms are strikingly correlated, but the progress also

appears to have been essentially parallel. This strongly suggests that there was a single grammatical development regulating the spread of both forms.

The first step in building the analysis of this grammatical development is to note (following Lightfoot (1981)) that the innovation of indirect passives required a change in the abstract Case system: as long as indirect objects were inherently Case marked (for oblique case), they couldn't be subjects of passives without violating the Case filter; or conversely, once indirect passives began to appear, it must have been the case that they were no longer inherently Case marked. As a tentative characterization of the period in question, I suggest that the mechanisms of a grammar which block indirect passives also allow verb movement, and thus there was no need to use 'do' in order to allow the verb to stay *in situ*; however, the change in the Case system which allows for the innovation of indirect passives also affects the way internal arguments are checked for Case, and thus the use of 'do' allowed the verb to stay *in situ* for Case reasons.

2.2 'That-trace' Effects

Lujan (1994) provides an independent argument for the idea that there is a relationship between Case and verb movement. She presents an analysis for various differences between Spanish and English which is based on the interaction of verb movement and the assignment of Case. Her proposal requires modifications of certain standard assumptions concerning Case assignment and feature percolation, but the range of phenomenon she addresses, from sentential complementation to obviation and locality, suggests that there are interesting empirical gains to be made if we view Case from her perspective.

The basic proposal is that in a language with verb movement such as Spanish, Case is assigned to the head of the complement, while in English, a language without verb movement, Case is assigned to the Spec of the complement. While Lujan does not provide a principled account for why grammars should work in this way, her basic assumptions allow for an account of why 'that' can be deleted in English and why 'que' cannot be deleted in Spanish. The explanation is that the English Case system does not directly Case mark the head of the complement, 'that', and so it can be deleted; in Spanish, on the other hand, 'que', being the head of the complement in a language with verb movement, is directly Case marked, and thus must receive phonetic interpretation.¹

The analysis extends to 'that-trace' effects in the following way. If we assume that the Spec CP position of an embedded clause is Case marked by the matrix verb, as it would be in English, then movement of a Nominative subject through the embedded Spec CP creates a chain with contradictory Case marking: Nominative from the embedded subject and Objective from the matrix verb. On the other hand, object extraction across 'that' is not blocked because there is no Case clash at the Spec CP node. Furthermore, extraction of an embedded subject across 'que' in Spanish does not create a Case conflict because C, not Spec CP, is Case marked by the matrix verb, and there is thus no Case clash.

Given this much, Lujan must make other adjustments to standard assumptions in order to account for why subject extraction structures in English are "saved" by the absence of 'that'; the idea she develops is that in the absence of a lexical complementizer, the embedded IP moves to CP in order to be licensed by a lexical head, i.e. the matrix verb. Under this view, she argues that the notion of "'that' deletion" is misguided: what is actually happening is that what we have usually assumed to be a "deleted" 'that' is actually never part of the derivation. In those instances when 'that' is not generated in C, then the

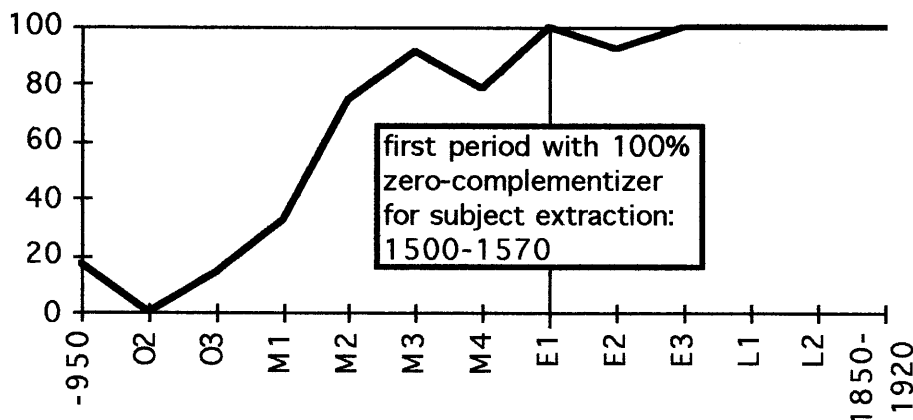
¹ Case marking of sentential complements is one of the additional assumptions Lujan makes.

embedded IP raises to CP in order to be licensed by a lexical head, and the abovementioned Case conflict is avoided.

While Lujan leaves open how certain technical details of her proposal works, there are nevertheless intriguing implications for the investigation being developed here. If Lujan's basic approach is on the right track, then it gives us a tool to use in the investigation of the relationship between Case and the loss of verb movement in English: Lujan's theory clearly predicts that there should be a chronological parallel between the innovation of 'that-trace' effects and the loss of verb movement. Furthermore, given the clear relation between the loss of verb movement and the use of 'do', the theory predicts that the development of 'that-trace' effects should coincide with the spread of 'do'. The predictions are borne out, as illustrated in a study by Bergh and Seppanen (B&S) (1992).

In a corpus of texts from 950-1920, B&S looked for instances of subject extraction with overt versus null complementizers. The results of their study are reproduced in Figure 2.²

Figure 2: Progress of 'that-trace' effects



The striking feature of the graph in Figure 2 is that the first period in which 'that-trace' effects are exhibited in 100% of the relevant 'that' deletion structures is the period 1500-1570. Recall the previous discussion concerning the use of 'do' in affirmative declaratives, namely that 'do' in affirmative declaratives decreased after the period 1550-1575; I have yet to make explicit the relation between 'do' in affirmative declaratives and the loss of verb movement, but assuming for the moment that the decrease of 'do' in affirmative declaratives indicates the loss of the productive use of verb movement, then the diachronic progress of 'that-trace' effects supports the predications made by Lujan's analysis of the relation between verb movement and 'that-trace' effects.

2.3 Summary

Discussing the spread of 'do' and the loss of verb movement in terms of extra-grammatical simplicity does not provide a way to address the range of innovations which developed simultaneously. The chronological parallels between the spread of 'do', the

² Figure 2 is produced by combining the information in B&S's Table 3, their Figure 1, and their discussion of Figure 1 on page 136. Unfortunately, they do not provide the breakdown of the data from which their Figure 1 was produced; also, the graph in their Figure 1 is not the clearest representation one might hope for. The graph in my Figure 2 is a good-faith attempt to accurately re-represent the information which B&S provide in their Table 3 and Figure 1.

innovation of indirect passives, the development of 'that-trace' effects, and the loss of verb movement suggest that these various changes were related to a single grammatical development. The obvious role of Case in the innovation of indirect passives, combined with Lujan's proposal concerning the relation between verb movement and Case, suggests that the grammatical change in question involved the abstract Case system.

Given the empirical paradigm at hand, there are (at least) two theoretical issues to be resolved: 1.) why should the decreased use of 'do' in affirmative declaratives be taken to be an indication of the loss of verb movement, and 2.) can the (early) increase in the use of 'do' be explained as a matter of economy of derivation, and if so, how (if at all) does Case fit into the picture? These issues are discussed respectively in the next two sections.

3. 'Do' as a First Resort

In order to interpret the decreased use of 'do' in affirmative declaratives as an indication of the loss of verb movement, I propose that we invert the common notion of 'do support' and assume that 'do' is a bound inflectional element which must be supported by an independently projected element of Infl --- either negation, emphasis, or (certain) features driving I-to-C movement.³ Furthermore, I must assume that derivations using 'do', if licensed, are more economical than derivations with affix hopping.

The theoretical motivation for this revised view of 'do' follows from the fact that the current notion of 'do' as a "last resort" cannot be extended to a diachronic account of the spread of 'do': if 'do' is a "last resort", and Middle English grammars allowed verb movement, then why should speakers have ever adopted the "last resort" option during the transition period? On the other hand, if we assume that 'do' is a "first resort" which can provide more economic derivations, then the diachronic development of 'do' is not a problem. Given the diachronic considerations, trying to maintain a "last resort" perspective on synchronic 'do' means that at some point 'do' would have changed from a "first resort" to a "last resort". Clearly the more attractive alternative is to regularize our perspective on 'do' and assume that its use both initially and currently is part of a "first resort" optimal derivation.

Under the assumption that 'do' is licensed by an independently projected element of Infl, we can consider in more detail the progress of the loss of verb movement. Roberts (1993), based on a cross-linguistic study of overt agreement paradigms and overt verb movement, finds that the minimal amount of agreement needed to trigger verb movement is equipotent markers for number.⁴ In other words, verb movement is overt in languages with (minimally) overt distinctions for number Agreement. He further recognizes the problem this poses for the history of English: the equipotent number distinction has more or less disappeared by 1500, but verb movement continues to be robust and productive for at least another 50-75 years.

In answer to this problem, Roberts assumes that, for some period, children acquiring Middle English had hypothesized a purely abstract Agr⁻¹ which forced the verb to move to Infl. Of course, a deep question is why children should have stopped

³ Sabine Iatridou (personal communication) points out, given counterfactuals like *Had John known . . .*, that it might be too general to say that any features driving I-to-C movement license 'do'; however, Mitchell (1985: 848) cites *Did he but know it, . . .* as a Modern English example of a conditional expressed through inversion. I will not explore this issue here except to note that the precise status of the features which license 'do' in C is not immediately obvious.

⁴ Roberts uses the Agr⁻¹ formalism, but we can just as easily say that there must be equipotent number markers for Agr to be strong.

hypothesizing the abstract Agr⁻¹ (and presumably the answer has to do with the decrease in the number of overt verb movement structures in their PLD), but the important point is that the abstract Agr⁻¹ continued to trigger verb movement. Now reconsider the “first resort” view of ‘do’: as long as grammars contained an Agr⁻¹ element, then ‘do’ would have been licensed in affirmative declaratives. However, once the abstract Agr⁻¹ was no longer hypothesized, then overt verb movement would no longer be triggered, nor would ‘do’ be licensed in (unemphatic) affirmative declaratives.

Under this analysis of ‘do’ and verb movement, the chronological correlation between the use of ‘do’ in affirmative declaratives and the beginning of obligatory ‘that-trace’ effects is not surprising: 1.) given Lujan’s proposal, ‘that-trace’ effects in English are an indirect consequence of the fact that English does not exhibit verb movement, and 2.) a “first resort” view of ‘do’ predicts that the use of ‘do’ in affirmative declaratives should decrease once verb movement is no longer a productive mechanism; therefore, what is otherwise a mysterious historical parallel, namely the pattern of ‘do’ in affirmative declaratives and the progress of ‘that-trace’ effects, is accounted for in this analysis by the mutual relation to the loss of verb movement.

Before moving on to the question of how the progress of ‘do’ can be accounted for in terms of Case and economy of derivations, there is a wrinkle in the “first resort” analysis of ‘do’ which is worth addressing. As stated, the prediction is that ‘do’ should have always been licensed in negative structures, but the patterns found by Ellegard show a downturn in the use of ‘do’ in negatives in the same period as the decrease in affirmative declaratives. If the synchronic account of ‘do’ is that the negative marker provides the necessary support for ‘do’, then something must be said about why the use of ‘do’ should have decreased at all in negatives.

The first line of explanation might come from a fact noted by Ellegard (1953:169): “negative ‘do’ holds its own better, and affirmative ‘do’ less well, in the colloquial group [of texts] than in the literary one” (my annotation --- MDA). Thus, the decrease might be an accident of the texts, and that in terms of colloquial use, ‘do’ was always equally licensed by the negative marker. However, a more satisfactory explanation can be seen if we consider the fact that the status of the negative marker was in flux during Middle English, namely it was changing from an adverb-like element to an inflectional head. Thus, if a structure contained the adverb-like negative, then ‘do’ would not have been licensed.

The subsequent upturn in the use of ‘do’ in negatives after 1600 suggests that the negative marker was moving towards its modern status as an inflectional head. Evidence which supports this view is seen in the fact that the use of ‘do’ in negative imperatives stayed very low until precisely the period when the use of ‘do’ in other negatives started its renewed increase; in other words, the sudden increase in the use of ‘do’ in negative imperatives corresponds with the renewed increase of ‘do’ in other negatives because the status of the negative marker had taken on its modern capacity to license ‘do’, and thus the previous imperative forms were no longer necessary (or perhaps even possible, given the loss of verb movement).

4. The Economy of ‘Do’

In the previous section, I simply assumed that the use of ‘do’, if licensed, provides a more economical derivation than affix hopping. Moreover, I suggested that viewing ‘do’ as a “first resort” would allow for a (somewhat) more straight-forward discussion of the diachronic process; however, even if we accept these assumptions, it is still unclear why the use of ‘do’ should have been more economical than verb movement (if it was), nor have we made any progress in explaining the patterns found by Ellegard. In order to make

some headway in this regard, I will first provide independent motivation for revising the minimalist framework in such a way that the diachronic problem of 'do' can be addressed. From there I will move on to showing how Case considerations, combined with the economy provided by 'do', can account for the patterns in the use of 'do' in different sentence types.

4.1 An Argument for a Non-lexical Array

Essentially, the problem is this: Chomsky's (1957) intuition that 'do' is introduced by rule in order to be the bearer of an unaffixed affix cannot be implemented in a system in which the computational system is fed by an array of fully specified lexical items.⁵ On the other hand, modifying the minimalist framework so that lexical insertion occurs at Spell-Out allows the system to "introduce" 'do' in the course of the derivation, and ultimately provides a means for discussing the relationship between the spread of 'do' and the loss of verb movement. The following discussion provides independent motivation for the proposal that the insertion of specific lexical material occurs at Spell-Out.

Consider the contrast in (1-2):

- (1) Joan is smarter than John.
- (2) *Joan is more smart than John.

If the initial numeration contains fully specified lexical items, as Chomsky (1993, 1994) proposes, then the contrast between (1) and (2) cannot be addressed: the initial arrays are distinct, so the derivations are not comparable, and thus there is no way for (1) to block (2). Furthermore, given that (3) is well formed:

- (3) Joan is more intelligent than John.

it is unlikely that (2) does not converge for some independent reason. If (2) converges, then it can only be ruled out by comparison to (1), but by definition, the derivations of (1) and (2) cannot be compared; the contrast in (1-2) is unexplained, indeed unaddressable.

However, if we assume that the numeration is an arrangement of purely abstract markers of lexical and functional heads, then the initial array for (1-2) is identical because they both contain the abstract marker for the head of a comparative phrase, call it [compar], and an Adj. The computational system manipulates these purely formal objects, and then specific lexical material is inserted at Spell-Out. Thus, assuming that both /-er/ and /more/ are possible Spell-Out objects for [compar], then we can posit that the derivation of (1) involves adjunction of the Adj to [compar], while the derivation of (2) does not. In other words, *smarter* is inserted at Spell-Out if the Adj has adjoined to [compar], but *more smart* is inserted if the Adj has not.

Given this notion of a non-lexical array, the derivations of (1) and (2) are comparable because they are computed based on the same numeration. Now that the derivations can be compared, we can allow the derivation of (1) to block (2) based on a principle of economy of insertion: inserting one word, when possible, is more economic than inserting two. Clearly this principle only comes into play in the specific instances where a synthetic form exists in the lexicon, as with *more smart* → *smarter*. As such, this

⁵ Of course we must transpose "introduce by rule" to current theoretical machinery, but the basic issue is whether a lexical item can be introduced in the course of a derivation; as Chomsky (1993, 1994) discusses it, there is no straight-forward way to introduce a lexical item in the course of a derivation.

approach provides an account for the general cross-linguistic tendency for synthetic forms to block periphrastic forms, as discussed by Poser (1992).

There is no doubt that the ramifications of viewing Spell-Out as lexical insertion are far-reaching; however, as illustrated above, there are empirical reasons to consider an alternative formulation of the numeration. Returning to the discussion of the spread of 'do', we see that the pattern of 'do' in different sentence types is another problem for a fully specified lexical array: without further emendation about how cooccurrence restrictions predetermine the mutual suitability of items in the array, there is no reason that the distribution of 'do' should not have been random across different sentence types and with different predicates. In other words, it is unclear why it should have been the case that the transitivity of the main predicate should predetermine the likelihood that 'do' was contained in the initial array; the same question arises for questions and negation.

For Modern English we can assume that if 'do' (or some other modal/auxiliary) is not generated in the array for questions or with negation, then the derivation will crash. For Middle English, on the other hand, since verb movement was an option in the grammar, then it is not so clear that there should be a necessary reason that the array leading to e.g. a negative question would also be the array which most frequently contained 'do'. On the other hand, if we approach the problem from the perspective that the array does not contain specific lexical material, then we are able to explore the possibility that 'do' tended to occur more frequently in specific sentence types, or with certain predicate types, because there was something about the derivations of those structures which lead to the insertion of 'do' rather than the overt movement of the verb.

4.2 Case and Verb Movement

Recall the basic proposal that the histories of 'do' and indirect passives are similar because of a changing Case system: whatever mechanisms of Case block indirect passives also allow verb movement, and when the Case system had begun to allow indirect passives, so the capacity for verb movement had begun to collapse. Furthermore, recall the tentative proposal that the use of 'do' in different sentence types can be accounted for in terms of the nature of the relationship between a (potentially moved) verb and its objects. The idea I will explore is that Case checking became dependent upon an LF incorporation process which, given economy considerations, effectively forced the verb to stay *in situ* until LF.

Let me point out now that there are two issues which I will not explore in detail. The first issue is how the distinction between structural and inherent Case will be represented in the minimalist framework: given that Case assignment is recast as feature checking under Spec-Head agreement, it is unclear how the structural/inherent distinction can be captured. My suggestion at this point is that structural Case is checked by a functional head while inherent Case is checked by a lexical head, but I will not develop any explicit representation.

The second issue is the nature of the relationship between theta roles and (inherent) Case. The issue becomes particularly messy when prepositions are considered, for even though prepositions are often assumed to be structural Case assigners, we must allow prepositions to mediate (if not determine) theta role interpretation based on contrasts like *I talked to John* versus *I talked about John*. The role of prepositions will be a factor in my discussion of the change in the Case system, but as my aim here is to provide an account for the spread of 'do' in terms of grammatical economy, I will remain agnostic as to whether prepositions 1.) merely mediate the semantic relationship between a predicate and its object, or 2.) determine the semantic interpretation of an object.

Returning now to the spread of 'do', consider the basic scenario (before verb movement was lost) given the notion of a non-lexical array: through the course of a derivation, the representation of the predicate might either move or stay *in situ*, and all other things being equal, the derivation which allows the use of a synthetic form should prevail.⁶ However, consider the ramifications if the Case checking features of the predicate (or the theta mediating features of a VP internal preposition,) require the LF incorporation of the preposition into the verb in order for the Case of the internal argument to be checked. If incorporation into traces is blocked (and I will argue momentarily for such a constraint), then the derivation in which the verb has overtly raised will either not be able to Case-check the internal argument, or extra grammatical machinery would need to be employed in order to create a representation in which the necessary incorporation process could happen.

Allow me to draw this out again. (Initially I adopt the notions of traces and reconstruction, but I will recast the story below within the copy theory of movement.) Suppose the representation of the verb moves prior to Spell-Out. Specific lexical material is inserted at Spell-Out, and so the verb is inserted somewhere in a projection well above VP. Now, in order for the derivation to lead to a fully interpretable object, the Case features of the internal argument must be checked before the representation reaches the LF interface. However, as there is only a trace of the verb inside the VP, a problem arises: a semantically functional preposition needs to incorporate into the verb, and since incorporating into the trace is not sufficient, one of two things can happen: either the derivation will crash, or reconstruction applies so that the preposition can incorporate into the verb. Reconstruction leads to a convergent derivation, but it also counts as a derivational step.

On the other hand, consider the ramifications of the availability of a periphrastic device, namely 'do', which checks the V-features in I and therefore allows a derivation to proceed without the verb being inserted into the phrase marker outside of the VP. In such a situation, the derivation which inserts the verb in the VP (and 'do' in Infl) allows the incorporation of the preposition into the verb to happen directly rather than requiring the further derivational step of reconstruction. Under this view, the use of 'do' did not increase simply because it allowed the verb to Procrastinate movement, but really it allowed a more economic derivation in terms of meeting Case requirements.

Given the generalized proposal, the patterns which Ellegard found can be accounted for in the following way. First, 'do' would have been more likely in questions because V-to-I-to-C movement creates a layered trace through which the reconstruction process would have had to apply twice, thus making the derivation of a question with overt verb movement even more costly given the extra application of reconstruction in order to check Case. Second, the fact that 'do' was used more frequently in negatives than declaratives falls out under the stipulation that reconstruction across an intervening head is more costly than reconstruction to an adjacent head.

As it stands, the proposal accounts for the spread of 'do', including the patterns found by Ellegard, by assuming that the insertion of 'do' into Infl (at Spell-Out) provided a more economic derivation due to the fact that the use of 'do' allowed the derivation to avoid the (costly) derivational step of reconstruction. Central to the analysis are the stipulations that there is no incorporation into traces, and that reconstruction across a head is more costly than reconstruction to an adjacent head. In fact, both stipulations follow from minimalist assumptions: given a principle of full interpretation and a copy theory of movement, the derivations in which incorporation applies to the "trace" of the verb encounter the following problems.

⁶ Procrastinate would not be violated because we are still considering the period when verb movement was productive, thus there must have been a strong feature in Infl (potentially) triggering verb movement.

First, given a copy theory of movement, if the LF representation uses the lower copy of the verb (the others having been deleted, presumably), then the Case features are checked and the object can receive a theta interpretation, but the Tense information of the clause is lost, thus yielding a semantically uninterpretable representation. On the other hand, if the higher copy is used, then the Tense information is interpreted, but the semantic role of the argument is not readable because the Case information pertinent for that object was deleted with the lower copy. It would seem that the only potential way to save the derivation would be to let the preposition incorporate to the lower copy and then “move” up the V-chain, thereby creating a single verbal complex with both Tense and Case information.

However, even if such a movement could be motivated, the derivation invoking such movement would end up being less economical than (and therefore ruled out by) the derivation in which the verb didn't move until after the preposition had been incorporated into it: if the preposition incorporates before (LF) movement then the V-chain automatically includes the necessary Case information; if incorporation waits until after (overt) movement, then either the representation will be uninterpretable, or some extra derivational process will need to apply in order to create an interpretable derivation.

As for the stipulation that reconstructing across a head is more costly than reconstructing to an adjacent head, the problem is similar to the one outlined above, except that the issue becomes one of defining locality in minimalist terms. Recall that the head-locality stipulation was invoked to explain why ‘do’ was used more in negatives than declaratives. Thus the issue for this particular case is the how the head of negation functions in the phrase marker, as the following suggests. For V-to-I movement, what is the mechanism which gets the verb “across” the head of negation given that a strict version of head-to-head movement would require the verb to at least “touch-down” at the head of negation? To deal with this, we must assume some type of minimality such that negation doesn't have the right kind of features to be “visible” for verb movement, and so the verb can cross it without violating head-to-head movement.

In terms of why “reconstruction” in negatives would be more costly than in declaratives, (i.e., over a head rather than to an adjacent head,) the problem becomes one of trying to apply the LF “movement” of the incorporated preposition “across” the head of negation: as before, if the preposition doesn't climb the V-chain before the interface, then the representation will fail to yield an interpretation; given that the preposition must climb in order for the derivation to be interpretable, then climbing “past” the head of negation will require some device to extend the domain of possible targets for movement analogous to the minimality condition which allows the verb to skip the head of negation; on the other hand, if there is no intervening head, then there is no need for the domain of possible targets to be extended, and the preposition can move to the verb (relatively) unencumbered.

One aspect of the discussion which is perhaps misleading is that it might seem as though I have assumed that the derivations of, say, negative questions could be compared to that of, say, affirmative questions, and that ‘do’ was used more in the former because of economy considerations. Of course this cannot be the case: the derivation of one type of sentence cannot be compared to the derivation of another type. The idea behind my proposal is to recognize that language change is not overnight, i.e., the change in the Case system would have evolved over time and varied with respect to given lexical items and given speakers. The patterns of ‘do’ fall out as a consequence of the overall products of a given grammar at a given time. In other words, for a given speaker at a given time, the accumulated effects of 1.) the changing Case system, and 2.) the (potential) derivational complexity of overtly moving the verb, meant that for negative questions, the system more frequently took advantage of the availability of periphrastic ‘do’. On the other hand, for

affirmative declaratives, given that the cost of moving the verb overtly would have been relatively insignificant, then the accumulated effects of the changing Case system would have taken longer to show significant influence in terms of the total number of times that the system used 'do' in the derivations of affirmative declaratives.

Clearly there are important issues involved in relativizing economic cost, but the basic intuition is that as the checking of Case features became more integrated with preposition incorporation, then 'do' came to be used more often as a way to allow the V head to remain *in situ* until LF. The question now, of course, has to do with the nature of the development of a Case system which invokes the incorporation of prepositions in order for the inherent semantic properties of predicates to receive interpretations. One historical development which suggests that the proposal is worth investigating more deeply is the fact that the number and use of prepositions increased dramatically in the progress of Late Old English and Early Middle English.

Superficially we can say that the increase in the number and use of prepositions was a consequence of the collapse of the morphological case paradigm, but given the (poorly understood) relationship between morphological case, abstract Case, prepositions and theta roles, the idea that the Case system changed in such a way that prepositions contribute to the checking of Case features, perhaps as a function of the assignment or transmission of theta roles, is not unreasonable. Whether incorporation is the right way to view the problem is not certain, but independent support for the idea can be seen in Lightfoot's (1991) analysis of the development of new accusative subjects and passives. Specifically, he proposes that the infinitival marker "to" came to coalesce with a governing verb and transmitted . . . both head-government and case" (Lightfoot (1991: 89-90)).

The head-government properties which Lightfoot invokes will not be (directly) pertinent given a minimalist framework, but given the role of Case in his analysis, and given that the innovations he discusses all occur in precisely the same period in which 'do' was replacing verb movement, then extending his "coalescence" proposal to include the incorporation of prepositions allows for an account of the spread of 'do', as argued above. Note also that the Hornstein and Weinberg's (1981) "reanalysis" analysis of English preposition stranding is exceedingly similar to the proposal developed here; in both analyses there is the creation of a verbal complex motivated by Case considerations and involving prepositions. Thus, both Lightfoot's coalescence process and Hornstein and Weinberg's reanalysis rule provide independent motivation for the proposal that the English Case system underwent a change which involved the incorporation of prepositions into a verb in order for (inherent) Case to be checked.

5. Summary

The central aim of this paper has been to show that the spread of 'do' in English, as well as the related loss of verb movement, is best accounted for by invoking a notion of simplicity in the grammar, not in acquisition, nor in processing. Only by looking for a singular adjustment in the grammatical system can we account for the fact that the innovation and spread of periphrastic 'do' so closely aligns with the innovation and spread of indirect passives. Additionally, the concurrent development of 'that-trace' effects suggests that abstract Case and verb movement are not (necessarily) absolutely independent mechanisms. An account for the chronological similarity of these various changes follows from a theory in which derivations became more economical when the verb remained *in situ*, not because of Procrastinate, but because Case checking would have required additional derivation complexity had the verb moved overtly.

A second aspect of the discussion has been to provide motivation for an alternative view of the minimalist framework, namely that the initial numeration contains purely abstract markers for lexical and functional heads, and that the insertion of specific lexical material is a function of Spell-Out. Motivation for this view comes from both the diachronic study of 'do' as well as the synchronic phenomenon in which synthetic forms block periphrastic forms.

References

- Bergh, G. and A. Seppanen (1992), "Subject Extraction in English: The Use of the That-Complementizer", in *English Historical Linguistics 1992*, Fernandez, Fuster, and Calvo, eds., CILT 113. Philadelphia: Benjamins.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. (1993). "A Minimalist Program for Linguistic Theory", in *The View from Bldg 20*, Hale, K. and Keyser, S., eds., 1-52. Cambridge: MIT Press.
- Chomsky, N. (1994). *Bare Phrase Structure*. MIT Occasional Papers in Linguistics 5
- Clark, R. and I. Roberts (1993). "A Computational Model of Language Learnability and Language Change". *Linguistic Inquiry* 24.2: 299-345.
- Denison, D. (1993). *English Historical Syntax*. New York: Longman.
- Ellegard, A. (1953). *The Auxiliary 'Do'*. Stockholm: Almqvist and Wiksell.
- Hornstein, N. and A. Weinberg (1981). "Case Theory and Preposition Stranding". *Linguistic Inquiry* 12.1: 55-91.
- Kroch, A. (1989). "Function and Grammar in the History of English", in *Language Change and Variation*, CILT 52, Fasold and Schiffrin, eds., 133-172. Philadelphia: Benjamins.
- Lightfoot, D. (1981). "History of Noun Phrase Movement", in *Logical Problem of Language Acquisition*, Baker, C. and McCarthy, J., eds., 86-119. Boston: MIT Press.
- Lightfoot, D. (1991). *How to Set Parameters: Arguments from Language Change*. Cambridge: MIT Press.
- Lujan, M. (1994), "Case Marking, Verb Movement and CP Complements", in *Probus, Proceedings of 1 Mesa Redonda de Linguística Española, Universidad Autónoma Metropolitana-Iztapalapa, Mar. 24-26, 1994*.
- Poser, W. (1992). "Blocking of Phrasal Constructions by Lexical Heads", in *Lexical Matters*, CLSI 24, Sag and Szabolcsi, eds. Stanford University Press.
- Roberts, I. (1993). *Verbs and Diachronic Syntax*. Boston: Kluwer.

Department of Linguistics
1401 Marie Mount Hall
University of Maryland
College Park, MD 20742

mdarnold@wam.umd.edu