BENJAMIN SLADE

HOW TO RANK CONSTRAINTS

Constraint conflict, grammatical competition, and
the rise of periphrastic 'do'

Abstract. This chapter illustrates the virtues of the Optimality-Theoretic framework (Prince & Smolensky 1993) in explicating the course of syntactic change. The rise of *do*-support, a well-known change in the history of English, is taken as a case study. We investigate the patterns of variation inherent in linguistic change that occur as innovating forms replace conservative forms. We take the position that these periods of variation reflect competition between grammatically incompatible structures, i.e. conceptualizing variation and change in the surface structures of language as a reflection of alternation of different underlying grammars (Kroch 1989a,b; 1994), which themselves result from reanalysis by language learners (cf. Lightfoot 1991 et seq.). We argue that the notion of constraint competition inherent in Optimality Theory is advantageous in understanding language change as competition between contradictory grammatical systems. Also, we demonstrate the capacity of Optimality Theory as a means of describing systematic, grammatically-structured long-term linguistic change – particularly changes following an 'S'-curve pattern of linguistic renewal – as resulting from systematic re-ordering of precedence relationships amongst conflicting universal grammatical principles.

Keywords: Language change, language variation, Optimality Theory, periphrastic *do*, *do*-support, grammatical competition, partial constraint ordering, Constant Rate Effect, English modals, Middle English, Early Modern English, Present Day English.

0. INTRODUCTION: OPTIMALITY AND CHANGE[1]

Optimality Theory (Prince & Smolensky 1993) began to be employed in the examination of sound change soon after its inception (e.g. Jacobs 1995, Zubritskaya 1995, Bermúdez-Otero 1996). Under the Optimality-Theoretic framework, "What [grammars of different languages] share are the universal constraints and the definition of which forms compete; they differ in how the constraints are ranked, and, therefore, in which constraints take priority when conflicts arise among them" (Prince & Smolensky 1997:1605). From this definition, it is a logical step to conceptualize language change as *constraint reranking*. Our primary focus in this chapter shall be to argue that Optimality Theory provides a framework that allows for a description of grammatically-structured long-term changes (in the sense of Warner 1997), wherein the fine details of the pattern of a change, like that of the establishment of periphrastic *do*, can be understood as contingent, though not inexorably connected.

There has been a relative paucity of Optimality-Theoretic analyses of syntactic change (the exception being Vincent 2000), despite the opportunities afforded by the framework of Optimality Theory (OT) and the prominence of diachronic *syntax* within parametric generative grammar (Government and Binding / Minimalist

337

Program), such as Lightfoot (1979 et seq.), Kroch (1989 et seq.), Roberts (1985 et seq.), etc. This chapter, along with that of Larry LaFond, seeks to address this lack. We present an OT account of one of the most extensively studied syntactic changes: the rise and regulation of *do*-support in English (Engblom 1938, Ellegård 1953, Lightfoot 1979:ch. 2, Denison 1985, 1993, Kroch 1989a,b, Stein 1990, Rissanen 1991, Roberts 1993, Warner 1993:ch. 9, Garrett 1998, Nurmi 1999, Han & Kroch 2000, and others). In addition, we hope to shed light on the nature of the transitional phase during which speakers employ multiple constraint rankings, generating alternative syntactic constructions, following in the basic methodology of Kroch (1989a,b).

Optimality Theory is a formal theory in which the Language Faculty is characterized as a set of violable constraints whose interaction governs a structural input-output mapping. As such, OT is not specific to any particular component of grammar, and the formalism – broadly adopted in phonology – has been extended to syntactic analysis (e.g. Grimshaw 1993, 1997; Legendre 2001). Under this formalism, certain parallelisms between phonology and syntax emerge. For example, the occurrence of epenthetic segments in phonology in some sense parallels the use of expletive or 'dummy' elements in syntax, as neither epenthetic segments nor expletive elements are part of the input, but occur only when their presence is required by some other principle of the grammar. Epenthetic segments occur in order to satisfy 'Markedness' constraints on syllabic structure, such as the requirement that all syllables have an onset. However, there also exist 'Faithfulness' constraints, which require identity between an underlying representation and the output, or surface form. One Faithfulness constraint, FILL (so called because it requires that structural positions be filled with underlying segments), which prohibits the appearance of segments in the output that have no correspondent in the input, is violated by use of an epenthetic segment. From this point of view, one way in which Arabic and English differ is in the relative ranking of these two constraints, FILL and ONSET (see Tableau 1). In Arabic, ONSET takes priority over FILL; therefore, epenthesis is required if a syllable would otherwise lack an onset. English displays the opposite ranking, thus epenthesis is not employed even if this occasions an onsetless syllable.

Tableau 1. Classical Arabic (Prince & Smolensky 1993:24-27) vs. English syllabic structure

Arabic: /al-qalamu/ 'the pen'	ONSET	FILL
.al.qa.la.mu.	*!	
☞ .ʔal.qa.la.mu.		*
English: /ink/ 'ink'	**FILL**	**ONSET**
☞ . ink.		*
.ʔ ink.	*!	

Just as Arabic and English differ in their requirements on syllable structure, Italian and English differ in their requirements on clause structure (Tableau 2). In Italian, a verb lacking a semantic agent, such as the 'weather'-verb *piovere* 'to rain', faithfully surfaces without a syntactic subject. English, on the other hand, requires a clause to have a subject, even when this means using a semantically empty expletive such as *it* as in *It's raining*. Again, we may posit a constraint FILL,[2] which prohibits the appearance of elements in the output that have no correspondent in the input, i.e. prohibiting elements, like the *it* of *it's raining*, that lack semantic-content. Like the constraint in UG requiring syllables to have onsets, UG also contains a constraint SUBJECT,[3] which requires clauses to have subjects. Italian and English show the opposite rankings of these two constraints, as Tableau 2 illustrates.

Tableau 2. Italian vs. English weather-verbs

Italian: Piove 'rains'	FILL	SUBJECT
☞ Piove.		*
Ciò piove.	*!	
English: Rains 'rains'	SUBJECT	FILL
Rains.	*!	
☞ It rains.		*

In this chapter, we utilize the interplay of Markedness and Faithfulness constraints in OT to analyze the rise of *do*-support in the history of English. One of the features of OT that allows for a coherent description of *grammatically*-structured long-terms changes is that the principles of an OT interact in a vigorous way, unlike the independently-set parameters of Principles & Parameters. The chapter is organized as follows:

In §1, we develop an account of *do*-support in contemporary English in terms of a ranking of constraints and its corresponding harmonic ordering of candidates, utilizing, for the most part, a set of constraints developed in the OT syntax literature (e.g. Grimshaw 1997, Vikner 2001).

§2 investigates the evolution of linguistic innovations, such as *do*-support, focusing particularly on the variation present between the onset of a change – when an innovation, such as periphrastic *do*, first begins to substitute for a previously established structure – and the completion of the change, when the conservative form becomes obsolete. We will develop a formal means to express such variation within OT.

In §3 we trace the various linguistic reanalyses, and the resulting constraint rerankings, that successively alter the distribution of *do*-support from its first appearance to its categorical present-day English patterning. The proposed account links the development of the class of 'modal' verbs with the rise of periphrastic *do* – again, not inexorably, but in the sense that the development of *do* is partially contingent upon the creation of the modal class (and the introduction of *to* as an unbound infinitival inflection). Tracing the evolution of *do*, it is demonstrated that

the relative proportions of *do*-use across syntactic contexts are derivable to a significant extent from the interaction of the constraints proposed herein, and that the overall pattern of changes in the use of *do* cross-contextually largely follows from this interaction of constraints. Finally, we argue that OT offers a means of conceptualizing such a change as a coherent process, rather than a series of logically unrelated events.

§4 provides a summary and concluding remarks.

1. *DO*-SUPPORT IN STANDARD PRESENT DAY ENGLISH

We begin by considering the patterning of *do*-support in contemporary English. The following section presents a provisional OT account of *do*-support in Standard Present Day English [SPDE], to be slightly revised in §3. While the proposed account is novel, it incorporates many elements of earlier analyses; within OT, Grimshaw (1997), Vikner (2001); within the 'Principles & Parameters' [P&P] framework, Lightfoot (1979, 1991), Roberts (1993), Han & Kroch (2000), Han (2000); and in Head-Driven Phrase Structure Grammar, Warner (1993).

Quirk et al. (1972:79) give the following list of constructions in which the use of auxiliary *do* is obligatory in SPDE [for "simple present, simple past" read "in the absence of modals (*can, could, will, would, may, might, must,* etc.) and the auxiliaries *have, be*"]:

1. In sentences negated by *not* where the verb is imperative [(1a)], simple present, or simple past [(1b)].

2. In questions involving inversion where the verb is in the simple present or past tense [(2a)]; exceptions: positive *wh*-questions beginning with the subject [(2b)] and *yes-no* questions without inversion.

3. In tag questions [(3a)] and substitute clauses [(3b)] where the verb is simple present or past tense.

4. In emphatic or persuasive constructions where the verb is simple present, simple past [(4a)] or imperative [(4b)].

5. In sentences with inversion caused by certain introductory words such as the negative adverbs never, hardly, etc. when the verb is in the simple present or past tense [(5a,b)].

(1)-(5) illustrate the various cases given in Quirk et al. above:

(1)   a.   Don't write plays!
        a'. *Write not of tragedy!
        b.   Will doesn't write plays.
        b'.   *Will writes not plays.

(2)   a.   What did Will write?
        a'. *What wrote Will?
        b.   *Who did write plays?
        b'. Who wrote plays?

(3)   a. Did Francis write plays? He didn't, did he?
       b. Speaker 1: But you write plays.
           Speaker 2: Yes, so I do.

(4)   a. But Will did write plays.
       b. Please do write a play.

(5)   a. ....nor did Will write plays.
       b. Never did Will finish one play, but he began another.

We shall be concerned primarily with the constructions in (1b) and (2a), that is to say, *do*-support in negatives and interrogatives. As mentioned above, *do* never occurs in the presence of modals or auxiliaries. This is illustrated in (6):

(6)   a.   William can't write plays.
       a'. *William doesn't can write plays.

       b.   William hasn't written any plays.
       b'. *William doesn't have written any plays.

After discussing our basic assumptions concerning phrase structure in 2.1, we present an OT analysis of the syntax of matrix affirmative declaratives, negative declaratives and interrogatives in SPDE.

1.1. Phrase Structure

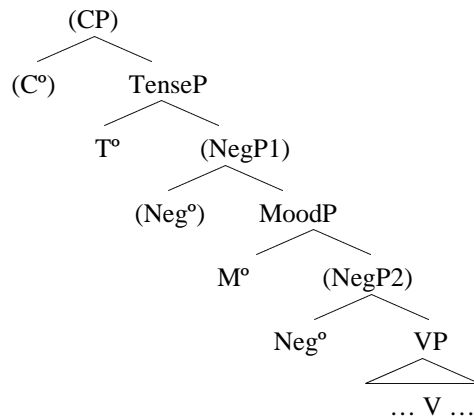We assume the following basic phrase structure in all syntactic analyses given herein:

Figure 1. Phrase-structure assumed herein.

As Figure 1 indicates, we posit two possible positions for sentential negation in English (cf. Han 2000). These different positions available for negation may perhaps reflect different scope-interpretation of the negation (see Zanuttini 1997, Cinque 1999).[4] However, this fact is not directly relevant to our investigation. For our purposes, it is sufficient that these two positions exist, regardless of their exact interpretation. In SPDE, a number of alternations support the availability of multiple positions for negation within the clause. The first is the different positions of negation in indicative clauses, as in (7):

(7)   a. The cat has <u>not</u> actually gone to London to see the Queen.
       b. The cat has actually <u>not</u> gone to London to see the Queen.

If Cinque (1999) is correct that adverbs are tied to invariant positions, then the alternation shown in (7) must be due to availability of different positions for *not*. Further, the variation in the positioning of negation in infinitives provides independent evidence for this conclusion.

(8)   a. For Harry <u>not</u> to have accepted the knighthood was foolish.
       b. For Harry to <u>not</u> have accepted the knighthood was foolish.

Other evidence for split IPs includes the positions of stranded quantifiers (Sportiche 1988), cf. Pollock (1989), etc. The division of IP into TenseP and MoodP is supported by examples such as those in (9) as pointed out in Baltin (1993), who draws attention to the asymmetry in the patterning of negation with respect to the infinitival marker *to* and the modals:

(9)   a.   Cædmon will <u>not</u> sing.
       a'. *Cædmon <u>not</u> will sing.

       b.   For Cædmon to not sing is a pity.
       b'. For Cædmon <u>not</u> to sing is a pity.

Note that *to* and modals do not display the same positioning with respect to negation, as the modals cannot occur linearly following *not*, unlike *to*, suggesting that *to* and the modals may occupy different structural positions in the syntactic tree. Further, the inability of the modals to occur linearly following *not* provides evidence that *to* occupies a position lower than that of modals. Thus we posit that *to* is generated in (and remains in) Mood°. We also posit that the modals (and *do*) are generated in Mood°, but that they raise to T°.

In the discussions to follow, we shall use 'derivation' terminology such as *movement*, for ease of conceptual exposition. This should not be taken to imply an actual derivation analysis. On the contrary, we assume a 'representational' outlook (or 'global' optimization), in which, for instance, traces represent dependencies rather than traces of actual 'movement' (for a true derivation approach to syntax

within the OT framework see the 'local' optimization analysis of Broekhuis & Dekkers 2000, Heck & Müller 2000).

In summary, we posit an IP split into TenseP and MoodP and the availability of two position of negation within the clause, one dominating VP and the other dominating MoodP. The following notation conventions are employed in the remainder of this chapter:

> **bolding** = raised **elements** and traces (**t**) of raising
> *italics* = lowered *affixes* and their traces (*t*)
> <u>underlining</u> = <u>empty heads</u>, containing neither overt elements nor traces of overt elements

1.2. Basic pattern of do-*support and verb-movement in SPDE*

We begin our analysis with an investigation of *do*-support in SPDE. This section presents a provisional account of *do*-support, which we shall revise slightly when we consider *do*-support diachronically in §3. Following Grimshaw (1997:375-376), we assume the input to GEN to be "a lexical head plus its argument structure and an assignment of lexical heads to its arguments, plus a specification of the associated tense and aspect". However, unlike Grimshaw, we assume that all matrix clauses invariably contain at least a certain number of functional projections, namely TenseP and MoodP.

The first context considered is the interrogative. *Do* appears in non-subject *wh*- and *yes/no* questions involving subject-verb inversion (10a). In contrast, in French, no 'dummy' element is employed (on the 'traditional' assumption that the pronoun *il* is in SpecIP); instead the main verb undergoes subject-verb inversion (10c), an option ungrammatical in SPDE (10b):

(10) a. Does he write plays?

      b. *Writes he plays?

      c. Écrit-il   des pièces dramatiques?
          writes-he the dramas

This suggests that French allows movement of lexical verbs like *écrire* 'write', whilst English does not, at the expense of inserting the semantically empty periphrastic *do*. Assuming that interrogatives quite generally require an inflected verb in C°, we posit two constraints to account for the noted crosslinguistic variation: the first, familiar from §0, is FILL, which penalizes the insertion of *do,* which is semantically empty (*pace* Tobin 2000):

> **FILL** (Prince & Smolensky 1993): An element in the output must correspond to an element in the input (i.e. no insertion of expletive or 'dummy'-elements, cf. FULLINT, Grimshaw 1997).

The second penalizes movement of verbs:

*HEAD-MOVEMENT [*X°MVT] (cf. NOLEXHDMVT, Grimshaw 1997):
Do not create a trace co-indexed with a verbal head (i.e. economy of
movement). [part of a 'family' of economy constraints, e.g. STAY]

The use of *do* in SPDE indicates that FILL is violated in order to avoid moving the
main verb, that is, to avoid additional violations of *X°MVT. Therefore *X°MVT
dominates FILL, as shown in Tableau 3:[5]

Tableau 3. do-*support vs. verb-raising in interrogatives*

	*X°MVT	FILL
☞ a. [_{CP} **Does** [_{TP} he **t** [_{MP} **t** [_{VP} write plays	**	*
b. [_{CP} **Writes** [_{TP} he **t** [_{MP} **t** [_{VP} **t** plays	***!	

Hence *do* is inserted in order to avoid movement of the main verb.[6] As for the factor
motivating the movement of elements to C°, a Minimalist analysis accounts for both
the fronting of *wh*-phrases to SpecCP and the filling of the head of CP by positing
strong features that require checking before LF. But we will follow Grimshaw
(1997) who proposes instead that *wh*-fronting occurs (not because of the need to
check a strong [+wh] feature in C°) to satisfy a high-ranked constraint that requires
that *wh*-phrases take scope over the clause:[7]

OPERATOR HAS SCOPE [OP-SC] (Grimshaw 1997): An operator must
take scope over the clause.

This constraint requires that overt operators, such as *wh*-words, and the covert
operators, such as that involved in *yes/no* questions, occupy a syntactic position
from which they take scope over the entire clause.[8] In a clause with a non-subject
wh-word this requires that the operator occupy a position dominating TP, thus a CP
is created and the operator moves into SpecCP, from whence it takes scope over the
clause.
    The CP, created for purposes of providing the operator with a scope position,
initially lacks a head. The movement of *do* or other auxiliary element into C°
supplies CP with a head. We follow Grimshaw in positing that CP, created solely for
reasons of scope, does not bear any 'strong features', but rather suppose quite
simply that the grammar requires that a projection not lack a head:

OBLIGATORY HEAD [OB-HD] (cf. Grimshaw 1997): A projection has an
overt head, or a dependency co-indexed with an overt head.[9]

The addition of these two constraints allows us to further explain the syntax of
interrogatives:

Tableau 4. do-*support, verb-raising and affix-lowering in interrogatives*

	Op-Sc	ObHd	*X°Mvt	Fill
☞ a. [CP What **does** [TP he **t** [MP **t** [VP write			**	*
b. [CP What **writes** [TP he **t** [MP **t** [VP **t**			***!	
c. [TP He **writes** [MP **t** [VP **t** what	*!			
d. [CP What <u>e</u> [TP he **writes** [MP **t** [VP **t**		*!		

Candidate 4c leaves the *wh*-word *in situ*, violating high-ranked Op-Sc. Candidate 4d, in which *what* is raised to SpecCP, but C° is left unfilled, is ruled non-optimal through its violation of ObHd. As in Tableau 3, candidate 4a is optimal, as the violation of Fill, inserting the semantically empty element *do* into Mood°, avoids a violation of *X°Mvt.

However, this set of constraints is insufficient to account for the syntax of affirmative declaratives (11a), which are ungrammatical with *do* in the simple past or present outside of emphatic contexts (11b).

(11)   a. Will actually writes plays, not screenplays.

       b. *Will does actually write plays....

       c. *Will writes actually plays....

       d. Will écrit   vraiment des pièces dramatiques....
          W.   writes truly   the dramas

The patterning of adverbs with respect to the main verb, e.g. (11a) vs. (11c), provides evidence that main verbs remain *in situ* in SPDE, compare with French (11d). This suggests that in such contexts, the affixes generated within Mood° and Tense° lower onto the main verb. Affix-lowering creates an improper dependency-chain in which traces c-command their binder, violating:

PROPER BINDING [PRBD] (Vikner 2001): In a dependency chain, X_i....Y_i....etc., co-indexed traces must *not* c-command an overt co-indexed element, i.e. no affix-lowering.

PRBD is a binary constraint (like ONSET in phonology); that is, for an individual dependency-chain PRBD is either satisfied or it is not – it incurs no gradient violations, unlike *X°Mvt, for instance (see McCarthy 2002b for arguments against gradience in OT).

We establish the ranking of PRBD in Tableau 5:

Tableau 5. do-*support, verb-raising and affix-lowering in affirmative declaratives*

	Ob-Hd	*X°Mvt	Fill	PrBd
a. [TP He **does** [ADV actually] [MP **t** [VP write plays		*!	*	
b. [TP He **writes** [ADV actually] [MP **t** [VP **t** plays		*!*		
☞c. [TP He *t* [ADV actually] [MP *t* [VP *writes* plays				*
d. [TP He e̲ [ADV actually] [MP e̲ [VP write plays	*!*			

Candidate 5d, in which T° and Mood° are left empty, is ungrammatical due to its violations of ObHd. Additionally, note that this also leaves the subject-verb agreement unrealized, for affixes, by definition, are bound inflections that require hosting by a lexical element. Here, since *do*-insertion also involves movement of *do* from Mood° to T°, *do*-support is non-optimal. The optimal candidate 5c violates PrBd by lowering of affixes onto the verb. The violability of PrBd has no effect on interrogatives (as in Tableau 4), as the lack of an element in the head of CP will produce a violation of undominated Ob-Hd.

Negative declaratives, such as (12a), are like affirmative declaratives in disallowing verb-movement (12c) – again, compare with French (12d) in which verb-raising does apply. Negative declaratives, however, differ from affirmative declaratives in the ungrammaticality of affix-lowering in the presence of negation (12b).

(12)   a. Will does not write any plays.

        b. *Will not writes any plays.

        c. *Will writes not any plays.

        d. Will (ne) écrit   pas des pièces dramatiques.[10]
             W. (*ne*) writes not any dramas

Descriptively, we may say that the presence of negation appears to block at least affix-lowering, while verb-raising fails to apply even in the absence of negation in SPDE as we observed in Tableau 5. However, we propose that negation in fact interferes with the movement in general, blocking both verb-raising and affix-lowering – this supposition is borne out by evidence from the diachronic evolution of English, to be seen in §3. The constraint violated by movement over negation is:

HEAD MOVEMENT CONSTRAINT [HMC] (Vikner 2001): No X° may intervene in an X°-chain. That is, if an X° is c-commanded by an element of an X° chain and itself c-commands another element in the same X°-chain, then it is bound by the antecedent (the lexical element) of the chain.

In our investigation, Neg° is the relevant intervening head. Note that, in negative declaratives with the higher negation, raising, whether of the main verb or of *do*, and affix-lowering, all cross over Neg°. This observation, combined with the optimality of *do*-support in negative contexts, tells us two things. The first is that the HMC must dominate *X°Mvt, otherwise affix-lowering would remain optimal in negative contexts. The second is that the high ranking of the HMC is insufficient in itself, since the use of *do* still involves a violation of the HMC (as will shall see in EME, the higher rate of *do*-support in negative contexts strongly suggests that we must consider that raising over negation incurs a violation of some sort). Therefore, we propose that affix-lowering over negation incurs a more severe violation of the HMC than does head-raising over negation. If we adopt the proposal that improper chains such as those created by affix-lowering are 'rescued' at LF through the creation of a proper binding relation by the raising of the verb (Chomsky 1995:ch. 2), then we see that affix-lowering involves 'crossing over' Neg° twice: once by the lowering of the affix and once when the chain is 'rescued' at LF (based on a vaguely similar idea of Rizzi 1990a:22-24). Tableau 6 shows the evaluation of candidates in negative declaratives formed with the higher NegP.

Tableau 6. do-support, verb-raising and affix-lowering in (high NegP) negative declaratives

	OB-HD	HMC	*X°MVT	FILL	PRBD
☞ a. [TP He **does** [NEGP1 not [MP **t** [VP write plays		*	*	*	
b. [TP He **writes** [NEGP1 not [MP **t** [VP **t** plays		*	**!		
c. [TP He *t* [NEGP1 not [MP *t* [VP *writes* plays		**!			*
d. [TP He <u>e</u> [NEGP1 not [MP <u>e</u> [VP write plays	*!*				

Negatives formed with the lower negation show an identical outcome, but a different assignment of constraint violations, as shown in Tableau 7. Here, as *do* need not cross over Neg°; the use of *do* in fact avoids the violation of the HMC altogether:

Tableau 7. do-support, verb-raising and affix-lowering in (low NegP) negative declaratives

	OB-HD	HMC	*X°MVT	FILL	PRBD
☞a. [TP He **does** [MP **t** [NEGP2 not [VP write plays			*	*	
b. [TP He **writes** [MP **t** [NEGP2 not [VP **t** plays		*!	**		
c. [TP He *t* [MP *t* [NEGP2 not [VP *writes* plays		*!*			*
d. [TP He <u>e</u> [MP <u>e</u> [NEGP2 not [VP write plays	*!*				

Hence, in SPDE, the choice between higher and lower negation does not affect the choice between *do*-support, verb-raising, etc., but, as we shall see in Section 3, in earlier stages of English this choice does have repercussions.

Having given a basic OT analysis of *do*-support in SPDE, we are prepared to investigate the development of *do*-support in the history of English, in §3, after a more general discussion of language change in §2.

2. AN OPTIMALITY-THEORETIC VIEW ON VARIABILITY AND CHANGE

2.1. Optimality and diglossia

A speaker's linguistic competence embodies knowledge of which syntactic structures are grammatical in his or her language. In actual language use we observe linguistic variation both between speakers and within individual speakers. This variation has nothing to do with speech errors, memory limitations or other factors that are commonly considered part of linguistic performance rather than linguistic competence, but rather involves knowledge of variation of grammatical structures. The use of alternate structures forms part of a speaker's linguistic competence, or knowledge of his or her language, and includes knowledge of which linguistic forms one uses in formal and informal speech-registers. Uriel Weinreich and his associates observe:

> The association between structure and homogeneity is an illusion. Linguistic structure
> includes the orderly differentiation of speakers and styles through rules which govern
> variation in the speech community; native command of the language includes the
> control of such heterogeneous structures. (Weinreich et al. 1968:187-188)

This sort of structural variation is glaringly present in the stages that constitute the establishment of *do*-support in English. In Middle English [ME (1066-1476)] the structures discussed in §1 above were not grammatical. In interrogatives, in the absence of modals or auxiliaries, instead of *do*-support, the main verb of the clause would undergo subject-verb inversion, e.g. *What wrote William?* rather than *What did William write?* Likewise, in negative declaratives, we find the constructions of the type *William wrote not* rather than *William did not write*. Only the late 15th-century marks the appearance of constructions with *do*, such as those shown in §1. However, when these *do*-constructions appear, they do not immediately replace the older *do*-less forms. Instead we observe variation during the Early Modern English period [EME (1476-1700)] between the older constructions without *do* and the innovative constructions with *do*. In certain periods, one also finds constructions with *do* that are ungrammatical in SPDE outside of emphatic contexts. Figure 2 (next page) shows the frequency of *do*-forms in different syntactic contexts from late ME to the end of the EME period. The differences in the frequency of use of *do* between contexts shall be discussed below, as shall the trajectory of *do* in different contexts and periods. For the moment simply note the alternation between clauses with and without *do*-forms and the general trend of increased use of *do* across contexts (excepting affirmative declaratives, also discussed below).

Variation in EME is not simply between speakers, as we find alternation between otherwise parallel constructions with and without periphrastic *do* in the EME period within the writing of a single author even in a single composition. (13), representing

early EME, is drawn from a collection of riddles, whose formulaic nature makes the alternation particularly striking. (14) is taken from later EME:

(13)   *c.* 1511 EME: *The Demaundes Joyous*, Wardroper (ed.) (Gray 1985:369)

   a. Why come dogges so often to the churche?
       'Why do dogs come so often to the church?'

   b. Why <u>doth</u> an oxe or a cowe lye?
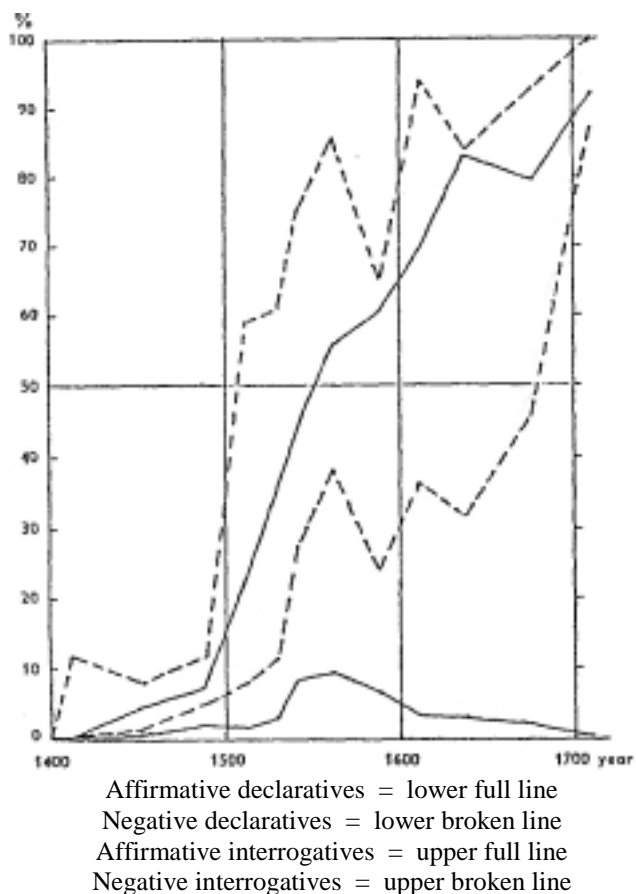   'Why <u>does</u> an ox or a cow lie down?'

Affirmative declaratives = lower full line
Negative declaratives = lower broken line
Affirmative interrogatives = upper full line
Negative interrogatives = upper broken line

Figure 2. Percent do-*forms: 1400-1700 C.E. (from Ellegård 1952:162).*

(14)   1605 EME: *The Advancement of Learning*, Francis Bacon

    a. So if any man think philosophy and universality to be idle studies, he <u>doth not consider</u> that all professions are from thence supplied and served. (2.0.8 (75))

    b. Out of which several inquiries there <u>do arise</u> three knowledges. (2.V.2 (100))

    c. I <u>doubt not</u> but it will easily appear to men of judgement. (2.VII.2 (106))

    d. This part of the metaphysic I <u>do not find</u> laboured and performed: whereat I <u>marvel not</u>. (2.VII.5 (111))

In (14d) we find in fact that both strategies are employed in juxtaposed clauses.

    In addition, the linguistic community of EME speakers, even within a particular geographic area, say the southeast, is not homogenous even in its variation, that is, different speakers use the alternate constructions at different frequencies. Robert Greene, Gabriel Harvey and Thomas Nashe were contemporary pamphleteers and even refer to (or, more specifically, attack) one another in their writings, and yet, we find variation in their respective uses of *do*-constructions. Our count of *do*-forms vs. non-*do* forms is shown in Table 1, from a study of a selection from each of these three authors, all three selections in fact being published in the same year, 1592. Two contexts are considered, affirmative and negative declaratives. Only examples in which *do*-support could occur are considered, i.e. non-indicative clauses are not counted nor are clauses containing modals or auxiliaries. Affirmative declaratives show alternation between forms with *do* and forms without. Negative declarative show three alternates: verb-*not*, *do-not*-verb and *not*-verb. Table 1 shows the fractions and corresponding percentages of occurrence of each type.

Table 1. do-*forms in pamphlet-writing, 1592*

	Aff. Decl.		Neg. Decl.		
	V only	***do*-V**	**V-not**	***do*-not-V**	**not-V**
R. Greene (Lamson & Smith 1942:428-433)	102/111	9/111	1/5	4/5	0/5
[from *The Repentance of Robert Greene, Master of Arts*]	**92%**	**8%**	**20%**	**80%**	**0%**
T. Nashe (Lamson & Smith 1942:446-453)	101/107	6/107	4/6	1/6	1/6
[from *Pierce Penniless, His Supplication to the Devil*]	**94%**	**6%**	**67%**	**17%**	**17%**
G. Harvey (Lamson & Smith 1942:436-445)	95/95	0/95	9/10	1/10	0/10
[from *Four Letters*]	**100%**	**0%**	**90%**	**10%**	**0%**

The samples examined are obviously very limited in scope, but still allow us to make some general observations. Most importantly, the complete lack of any occurrence of the type *do*-verb in the selection from Harvey contrasts sharply with patterns of the Greene and Nashe texts, which both show significant use of this type. Greene shows a high use of *do* periphrasis in both affirmative and negative contexts, whilst Harvey exhibits only one periphrastic *do* form in a negative declarative and Nashe stands somewhere in between as regards both affirmative and declarative contexts. This is consistent with Stein's observation that "the average figures of nine and three percent [of *do* in affirmative declaratives, between 1575 and 1625: see Figure 2] are quite misleading here, in that there are texts with hardly any *do* at all and texts with frequencies of around 50%....[the] author Boorde....has a *do* frequency of 64% in declarative sentences" (1990:109). Therefore, there exists rather wide divergence between individual texts, somewhat obscured by Ellegård's averaging.

Though its use seems more highly favored by some authors than others, a general description of the alternations in the EME corpus seems to be that *do* periphrasis occurs optionally. Contemporary linguistic theory – both in the OT and Minimalist formalisms – does not allow for the optional application of rules/operations or constraints. Part of the move away from optional rules is the recognition that many constructions that may appear to be the result of the optional application of a rule, e.g. active vs. passive constructions, are not actually semantically equivalent. In this vein, Legendre (2001) discusses the apparent variation in German between clauses like *Gestern wurde getanzt* [Yesterday was danced], *Es wurde gestern getanzt* [It was yesterday danced]: both meaning roughly, 'There was dancing yesterday'. Legendre shows that such constructions are inequivalent in logical form, as the seemingly free variants involve focusing and/or topicalization. She then presents a well-motivated OT analysis that correctly predicts the discourse contexts under which each of the constructions is grammatical. However, not all cases of 'optionality' are so easily circumvented. In particular, the variation between clauses with and without *do* in EME does not seem to involve any semantic differentiation.

The notion of co-existing grammars is what is required for an analysis of these remaining cases of optionality. As Lightfoot (1999:92) explains, "[A]pparent optionality [is] a function of coexisting grammars. Rather than allowing one grammar to generate forms *a* and *b* optionally, we would argue that a person has access to two grammars, one of which generates *a*, the other form *b*; the speaker has the option at any given time of using one or other of the grammars. This move reduces the class of available grammars, eliminating those with optional operations". The use of two (or more) grammars by a speaker is often termed *diglossia*, as distinct from bilingualism (Ferguson 1959, Lightfoot 1991:136-137). That is, 'diglossic' co-existing grammars are grammars that do not constitute different 'languages', i.e. are mutually intelligible, unlike 'bilingual' co-existing grammars.

Diglossia provides a useful formal conceptualization of cases of language change that occur via variation between alternate constructions, in which one construction is

increasingly substituted for by another. On this view, "change proceeds via competition between grammatically incompatible options that substitute for one another in usage" (Kroch 1994:180). Semantically-equivalent variants, within both OT and Minimalism, cannot be produced by the same grammar, and thus represent 'grammatically incompatible options'. The introduction of a linguistic innovation, its increasing substitution for a pre-existing formation, often leading to the loss of the latter, can be understood as the introduction of a new grammar generating the innovation, which, initially co-existing and competing with the older grammar, eventually replaces it, leading to the loss of the older form.

Within OT, the maintenance of multiple grammars by a speaker can be conceptualized by the formal device of 'floating' constraints (Zubritskaya 1995, 1997), *partial* constraint ranking (Anttila 1997b, Anttila & Cho 1998) or continuous ranking with noisy evaluation (Boersma 1997a et seq.), which allows one to describe variation as grammatical competition whilst capturing the continuity of the alternating grammars, i.e. the competing grammars are identical in many respects, differing only with respect to a few dominance relations. In standard OT, the set of ordered pairs of constraints, R, within the complete constraint set C has the properties of being irreflexive (no constraint dominates itself), asymmetric (if A » B, then *B » A), transitive (if A » B and B » C, then A » C) and connected (every constraint is ranked w.r.t. every other constraint). This defines a *total* order of constraints. However, the first three properties alone define a more general relation, which includes total orders as a special case: partial order. In a partial order, it is not a requirement that every constraint be ranked w.r.t. every other constraint, allowing for the possibility of indeterminacy in the grammar. We assume, as does Anttila (1997b), that while the constraint set C may be a partial order, on any given evaluation of competing candidates, the constraint ranking still need be a total order. Therefore, on an evaluation of candidates, constraints whose ranking is not completely determined 'fall into' one of their possible rankings. Thus, all actual productions still use a strict, total ranking of constraints to evaluate competitor forms, but which form is chosen may vary from production to production as the constraints governing the forms may vary in rank.

This OT-based definition of 'diglossia' is also more restrictive than simply allowing any set of co-existing grammars, in that only certain combinations of grammars define possible partial constraint orderings. E.g., where A,B,C are constraints, the set of grammars that contains the rankings A » B » C; B » A » C; and B » C » A is a possible outcome of a partial constraint ordering, as it requires only that B dominates C and that A is not ranked w.r.t. to B or C; but the set that contains A » B » C and B » C » A and *nothing else* is not consistent with a partial constraint ordering, as it would require that A » B and C » A be relations within R, where R is the set of ordered pairs defining pairwise relations between constraints – but as the relation B » C is also in R, transitivity requires that A » C also be a relation in R in addition to C » A, thus violating asymmetry.

In sum, the OT-based formalism for competing grammars provides a useful, restrictive definition of diglossia, which allows one to describe the variation inherent

in a language, dialect or even idiolect by means of 'multiple grammars' and at the same time to express the cohesion between these alternate grammars, as they overlap in many respects.

2.2. Evaluating patterns of change

 In order to understand change as a reflection of the competition of formal grammars as in §2.1 above, it is necessary to come to an understanding of the relation between a formal grammar of a language and the distribution of alternating (or 'competing') forms in that language. Some views on language change, such as Bailey (1973), characterize change in terms of the surface forms themselves, in terms of the lexicon. We, on the other hand, in the same 'generative' vein as Kroch (1989b, 1994), contend that changes in the distribution or frequency of surface forms, e.g. *do*, in a language are a manifestation of an underlying change in the grammar itself – not an 'analogical' spread of a lexical item. In this section we argue for the validity of this view.

Linguistic renewal often follows a trajectory that can be described by an 'S'-shaped curve, the *x*-axis representing time and the *y*-axis representing percentage of new forms, as shown in Figure 3. The curve is 'S'-shaped as innovative forms substitute for older forms very slowly at first, this replacement accelerating in the median stages before slowing again prior to the complete loss of the earlier forms.

> A given change begins quite gradually; after reaching a certain point (say, twenty per cent), it picks up momentum and proceeds at a much faster rate; and finally tails off slowly before reaching completion. The result is an ∫-curve: the statistical differences among isolects in the middle relative times of the change will be greater than the statistical differences among the early and late isolects.[11] (Bailey 1973:77)

The 'S'-curve is widely recognized in descriptions of change, cf. Weinreich et al. (1968), Kroch (1989a,b), Lightfoot (1991, 1999), Denison (1999), etc. The establishment of *do* in many of its contexts fits this model of an S-shaped curve, the abstract form of which is shown in Figure 3 below:
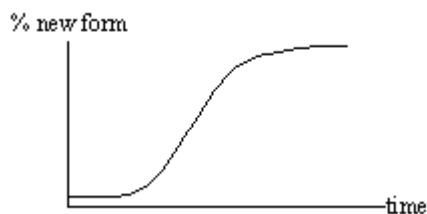

Figure 3. The 'S'-curve.

However, even a cursory examination of Figure 2 makes clear that the different syntactic contexts experience different relative proportions of the use of *do*-support throughout the EME period. In what way is the progression of *do*-support in each context related to the progression of *do*-support in the other contexts? We adopt the

model of historical change as the competition of grammatically incompatible options developed in Kroch (1989a,b; 1994), which crucially relies on the Constant Rate Hypothesis (Kroch 1989a,b). This hypothesis claims that, contrary to Bailey's assertion, the frequency of use of new forms in different contexts proceeds at the same rate in all contexts. A visual examination of Figure 2 makes it appear that the rate of increase is different for each context. This appearance, however, is deceptive in that it relies on the assumption that the average slope of the curves is a reasonable measure of the rate of increase, that is, that the curves are essential linear. However, the curves are actually best described by a logistic function and Kroch (1989b) demonstrates for several cases of change, including the rise of *do*-support, that when the curves reflecting the different contexts in which an innovation is present are fitted to this logistic function[12] (which approximates the shape of the 'S'-curve) taking into account differences in the values of the intercepts (difference in points of actuation or initial frequency between contexts), the slopes of the curves, that is to say the rates of progression in the use of *do*-support across contexts, are statistically *identical*. This is to say that the curves representing the progress of innovating forms in each context are identical, and simply displaced from one another either along the x-axis (if the initial frequencies are identical and the actuation sequential, see Figure 4) or the y-axis (if the actuation is simultaneous and the initial frequencies non-identical, see Figure 5).
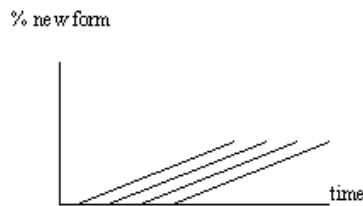

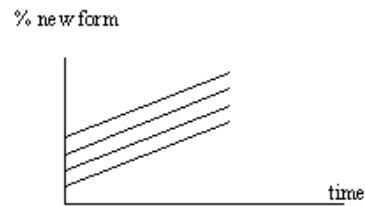*Figure 4. Serial actuation, equal rates of progression.*

Figure 5. Simultaneous actuation, equal rates of progression.

Kroch (1989a,b) shows that, having fitted the curves to the logistic function (see above), the increase in frequency of *do*-support advances at the same rate in the negative declarative, negative interrogative and affirmative interrogative contexts between 1400-1575 (Kroch 1989b:223-225). The rate of *do*-support in affirmative declaratives, however, does not show the same rate of increase as do the other contexts in the period considered. But using Ellegård's (1953:184) examination of the positioning of the adverb never with respect to clauses with and without *do* – e.g. *I saw never such a sight* (verb-raising), *I did never see such a sight* (*do*-support), *I never saw such a sight* (affix-lowering) – Kroch is able to make a rough estimate of the rate of non-application of verb-movement (after factoring out the estimated frequency of never occurring in a pre-INFL position) (Kroch 1989b:226-228). What he shows is that the combined frequency of the use of the strategies of *do*-support and affix-lowering in affirmative declaratives advances at the same rate as *do-*

support in other contexts. In other words, the option of verb-movement is what is being lost at the same rate in all contexts. The appearance of *do* is one result of the loss of verb-movement.

The Constant Rate Effect is consistent with the distribution of data in a number of changes, crucially including the rise of *do*-support, arguing against a conception whereby an innovation spreads from one context to another. The constant rate of progression of innovating forms, like *do*, across contexts reflects the fact that they all are surface manifestations of a *single underlying change* in the grammar. This is a crucial factor if one is to describe the course of a particular change in terms of the competition of grammars (see above 2.1).

Kroch (1989b) argues that the initial differentiation of the contexts in terms of frequency of use of *do* forms reflects "functional effects, discourse and processing, on the choices speakers make among the alternatives available to them in the language as they know it; and the strength of these effects remains constant as the change proceeds" (238). The Constant Rate Effect can only be mathematically proven to hold up to 1575, however, at which time, as an examination of Figure 2 reveals, we observe deviation from the general 'S'-curve pattern in many of the contexts. These deviations suggest that other changes begin to take place around this time, which distort the shape of the 'S'-curve. These points of deviation appear to affect contexts with respect to particular syntactic characteristics, e.g. the presence of negation – which indicates that contexts may also be directly distinguished by syntactic properties, rather than simply processing demands. If the contexts begin to show patterned divergence conditioned by syntactic properties in 1575, there seems no reason to suppose that the initial differences in frequency of *do* use may not likewise be directly due to their distinct syntactic properties. Our analysis of the rise of *do*-support, given in §3, considers that the differences in frequency of the use of *do* between contexts, as shown in Figure 2, indeed reflect the fact that the contexts are initially distinguished by syntactic properties, in a way that impacts the proportion of *do* forms that they exhibit.

The Constant Rate Effect supports a conceptualization of linguistic change couched in terms of competition of grammars. In the next section we present an analysis of the patterns of variation and change in EME clause structure utilizing the notion of partial constraint ordering or 'floating' constraints as the formal characterization of the competition of grammars, arguing that this extended OT formalism provides a more coherent framework for understanding the course of language change than does P&P.

3. THE RISE AND REGULATION OF *DO*-SUPPORT: AN OPTIMALITY-THEORETIC ANALYSIS

In Early Modern English *do* has not yet attained the categorical distribution we observe in Standard Present Day English. Ellegård's survey (1953) of the frequency of *do*-forms across a number of syntactic contexts from 1400-1700, based on a sample of over 10,000 tokens, summarized in Figure 2 above, shows the average

frequencies of *do* in affirmative and negative declaratives, affirmative and negative interrogatives and negative imperatives. One may observe that the plot of the frequency of *do* over time produces roughly 'S'-shaped curves for each of the contexts, as discussed previously. Note that before 1500 Ellegård gives figures only for texts that contained instances of periphrastic *do*: "It is obvious that the 15th century frequencies would become even smaller than they are if all the texts read had been allowed to contribute" (Ellegård 1953:159). Warner (1993:220) notes that "[Ellegård] used only 40 per cent of his data for the period 1425-75, and only 75 per cent for 1475-1500". Thus, the initial part of the 'S'-curve should in fact be thought of as being rather steeper than shown in Figure 2.

Representative tokens of sentences in the EME period with *do* are given in (15) and without *do* in (16), drawn from each of the syntactic contexts, excepting the negative imperatives (on which we do not focus in this investigation). The examples of the use of *do* in (15) include cases (15e), (15f), where *do* is ungrammatical in SPDE outside emphatic contexts. At the same time (16) shows sentences lacking *do*-support, in contexts where it is obligatory in SPDE.

(15)   a.   whose sore task **does not** divide the Sunday from the week...
           (*Hamlet*, II.i.83)

        b. **do** you **think** I am easier to be played on than a pipe? (ibid., III.ii.111)

        c. the mony is so little that it **doth not** suffice
           (Ellis, Orig. Lett. UU, cited Roberts 1993:251)

        d. ....that I **did** not believe...to be capable of...
           (William Aubrey, *Letter to John Dee*, 1577)

        e.   I **did** translate it my selfe into plaine English...
           (Roger Ascham, *The Scholemaster*)

        f. ....but verie warely **doeth** sell the same commoditie vnto vs.
           (Robert Hitchcock, *A Pollitique Platt for the honour of the Prince, the great profite of the publique state*, a3v)

(16)   a. I **loved** you **not**. (*Hamlet*, III.i.121)

        b. **Looks** it **not** like the king? (ibid., I.i.43)

        c. How **fares** my lord? (ibid., III.ii.255)

        d. What, noughty hoore, **caull** thou me goose steiler?
           (Agnes Wheitley, *Depositions and Other Ecclesiastical Proceedings from the Courts of Durham*, 1568)

        e. ....as the best of them disdained **not** the poorest....
           (Sir John Harington, preface to *Orlando Fuerioso*)

3.1. The development of periphrastic do *and unbound inflection:* to *and the modals*

The rise of periphrastic *do* is connected to other innovations of the late ME period: the development of the 'modal' class and the use of *to* as a marker of the infinitive. Both of these developments reflect the introduction of *unbound inflection* into the language, that is, an element that carries Tense, Agreement or Mood information or features and that does not form a morphological unit with the verb, being thus not 'bound' to the latter. *Do* is employed in superficially 'periphrastic' constructions prior to EME. However, its earlier uses in ME are 'contentful', unlike the modern semantically empty element. The exact semantics of the earlier 'periphrastic' construction is unclear. There is some evidence (Ellegård 1953, Denison 1985) that its ultimate source may lie in a causative construction parallel to that of SPDE *make*. Garrett (1998) also offers evidence that the periphrastic originates in a verbalizing *do*-construction, like that of SPDE 'He does windows' [= He washes windows]. By the late ME period, *do* appears to have been used as an aspectual element, marking perfect (Denison 1985) and/or habitual (Garrett 1998) aspect. Whatever the semantics of these earlier *do*-constructions, they are important because they established the superficial syntactic pattern of the empty-periphrastic: NP+*do*+'bare' infinitive VP ['bare'=without *to*]. In the late 15th-century, we observe the rise of the semantically empty periphrastic construction, evidenced by the increase in the use of *do* in all syntactic contexts.

The modern English modals, with the exception of *will* (which exhibits idiosyncratic irregularities, similar to those of the preterite-presents, eventuating in its becoming a part of the modal-class), are the descendants of a subset of the class of OE verbs sometimes known as the preterite-presents, as this class has present tense forms inflected like those of the strong past (apart from the present tense form –*st* of their 2nd person singulars). Thus, in the present tense, the 1st and 3rd person singular forms of this class (including *will*), are homophonous, distinguishing this class from other verb classes in OE. The OE preterite-presents semantically functioned primarily as ordinary verbs: OE *cann* 'to know, to be acquainted with (cf. Ger. *können*), 'to know how to' > SPDE *can*; OE *mæg* 'be strong, have power, be able' > SPDE *may*; OE *sceal* 'owe, must' > SPDE *shall*, etc. Though, as has been often noted, even in OE we can observe the occasional 'modal' (epistemic or deontic) use of the preterite-presents, especially *will* as representing future time (though often still with a tinge of the root sense of 'want'), these increase in frequency in ME.

Warner (1993) shows that the preterite-presents *gradually* developed many of their 'modal' characteristics, including the semantic shifts, inability to take direct objects, loss of finite forms, etc. However, around the beginning of the 16th century, we observe the loss of co-occurrence of 'modal' verbs from the majority of texts (the iteration of modals disappears completely from the standard dialects in the early 16th century (cf. Lightfoot 1979:ch. 2.1, Roberts 1993:294-295). Therefore, though the modal-class certainly experiences a certain amount of gradual development, particularly of the semantics of individual items, we maintain that the class undergoes a categorical syntactic reanalysis beginning in the late 15th century (cf.

Lightfoot 1979:ch. 2, 1991:140-143; see Plank 1984, Warner 1993 for critiques of Lightfoot's analysis). Additional evidence for this reanalysis, aside from the loss of iteration of modals, is that, as the use of *do* forms develops for lexical verbs in the 16th century, we never observe the pattern: periphrastic *do* + modal.

Another distinguishing characteristic of the preterite-presents is that throughout the history of English they almost invariably select the simple infinitive, with the suffix –*an*, rather than the complex infinitive, preceded by *to* and taking the suffix –*anne* (for the small set of exceptions to this rule of infinitival selection, see Mitchell 1985:§996 on OE, Warner 1993:138 on ME). However, the selection of a *to*-less infinitival complement was not as distinctive in OE as it is in SPDE. In general, OE verbs showed some variance in their choice of infinitive: some always took the complex infinitive, some always the simple, and some alternated between the two. The *to* of the complex infinitives seems to originally have been a preposition – the selection of prepositional infinitives is not unusual of Indo-European languages, cf. French *Il oublie d'aller à Paris* 'He forgets to go to Paris' [with a prepositional infinitive] vs. *Il espère aller à Paris* 'He hopes to go to Paris' [bare infinitival form]. As opposed to the Romance languages, however, in English the complex, prepositional infinitive steadily encroached on the simple infinitive, leaving the auxiliaries, modals and a small set of mainly causative [e.g. *make*] or 'sense-perception' [e.g. *see*] verbs as unique in selecting for a bare infinitival complement.

As noted, aside from these two classes, certain constructions involving *do* also exhibit this pattern, i.e. NP-*do*-'bare' infinitive complement, as remarked above. The sense of *do* in such constructions, before the rise of periphrastic *do*, has been variously analyzed as causative (e.g. Ellegård 1953) and/or aspectual, including either perfective aspect (Denison 1985) or habitual aspect (Garrett 1998). It is possible that *do* included all of these uses – the confusion between such meanings may in fact have played some role in producing the meaningless periphrastic. In any case, *do* is another member of the set of elements selecting for the simple infinitive – a class that becomes very restricted (to the modals, the sense-perception/causative verbs and *do*) during the ME period.

In order to analyze these developments, let us begin by attributing the lack of unbound inflection in the early ME period to a constraint penalizing its occurrence:

> **V+INFL**: A verbal head must be attached to Agreement, Tense and Mood features, i.e. no unbound inflectional morphemes.

In ME grammars, infinitival verbs raise to Mood° in order to pick up and host infinitival inflection (-*anne, an, -en, -e, -Ø*, depending on the period), at the cost of violating *X°MVT. This attests a grammar in which V+INFL dominates *X°MVT, as in Tableau 8 below.

Tableau 8. ME infinitives

	OBHD	HMC	PRBD	FILL	V+INFL	*X°MVT
☞a. ([PP to)[MP **V** [VP **t**						*
b. [MP to [VP **V**					*!	

Verb movement in infinitives in ME is attested by the pattern of negation in such constructs, in comparison to negative infinitives in EME and SPDE. We presume (17a) to be formed with higher negation, exhibiting the same sort of structure as EME and SPDE. In contrast, (17b) appears to show a construction with the lower negation, which in EME and SPDE would be equivalent to the type *to-not*-verb. The appearance of the verb before *not* in (17b) provides evidence that the infinitival verb in ME raises to Mood° to host inflection (at this stage –e).

(17)   a. ***not-to*-verb**
          that sche wuld   vwche-save
           that she would   promise
          <u>nowth   to   labowre</u>   agens   yw   in this matere....
          [NEGP1not   [PP/CP? to [MP **labour** [VP **t** against   you   in this matter....
                                        (*Paston Letters* 221.310 [cited in Han & Kroch 2000])

       b. ***to*-verb-*not***
          <u>to   spille   not</u>   oure tyme, be it short….
          [PP/CP? to [MP **waste** [NEGP2 not [VP **t**   our   time, be it short…
             (*Purvey's Prologue to the Bible* I,56.73 [cited in Han & Kroch 2000])

During ME, the ending of the complex infinitive weakened its inflection to –en. This ending in turn weakened to –e before disappearing about the same time as the -e(n) plural marking in East Midlands English, roughly at the beginning of the 16th century (Roberts 1993:261). This loss leaves the complex infinitives 'marked' only by prepositional *to*. We postulate that speakers reanalyze *to* as the marker of infinitival inflection – cf. Lightfoot (1979:186-199) and Roberts (1993:259-262), which means that *to* is now generated under Mood°, in the place of the formerly used suffix *–anne, –en, –e*.

    This reanalysis of *to*, as well as the reanalysis of the modals discussed below, are not simply reanalyses of the nature of certain elements nor are they solely changes in the lexicon – they are changes in the grammar. In OT, the lexical inventory of a language is a function of the grammar itself (the principle of 'Richness of the Base' – cf. Prince & Smolensky 1993:191ff), not a separate listing of elements, whether those elements be 'phonemes' or syntactic elements, like infinitival markers. Thus, when speakers reanalyze *to* as an inflectional element, they also posit a new grammar, in which an unbound inflection like *to* is grammatical, as shown in Tableau 9, where *X°MVT has been promoted above V+INFL.

Tableau 9. EME infinitives

	OBHD	HMC	PRBD	FILL	*X°MVT	V+INFL
a. ([PP to)[MP **V** [VP **t**					*!	
☞ b. [MP to [VP V						*

After this reanalysis, selection for the simple infinitive actually means selection of a 'bare' infinitive, i.e. an infinitive lacking overt infinitival marking (*to*). The modals, occurring before *to*-less infinitives, thus become ripe for reanalysis also, as they syntactically appear equivalent to *to*, that is, they pattern very much like inflectional elements. If modals are generated as unbound inflection, in Mood°, this reanalysis also accounts for their inability to co-occur, as multiple elements presumably cannot be generated under the same projection. Co-occurrence prior to the reanalysis would be possible as 'modals' were generated within potential recursive VPs. Thus by the same reanalysis by which *to* becomes an inflectional element, the modals attain their SPDE patterning, as shown in Tableaux 10 and 11.

Tableau 10. ME clauses with preterit present, or 'modal', verbs

	OBHD	HMC	PRBD	FILL	V+ INFL	*X° MVT
☞a. [TP NP **'modal'** [MP t [VP **t** [VP V						**
b. [TP NP **'modal'** [MP t [VP V					*!	*
c. [TP NP t [MP t [VP *'modal'* [VP V			*!			
d. [TP NP e [MP e [VP 'modal' [VP V	*!*					

In ME, 'modals' are generated under a VP as verbs, raising to T° as in Tableau 10. The reanalysis in EME leads to their generation under MP, saving a violation of *X°MVT, at the cost of violating V+INFL, as in Tableau 11.

Tableau 11. EME modals

	OBHD	HMC	PRBD	FILL	*X° MVT	V+ INFL
a. [TP NP **'modal'** [MP t [VP **t** [VP V					**!	
☞b. [TP NP **'modal'** [MP t [VP V					*	*
c. [TP NP t [MP t [VP *'modal'* [VP V			*!			
d. [TP NP e [MP e [VP 'modal' [VP V	*!*					

Turning back to our protagonist, periphrastic *do*, let us note the similarities between *do* and the modals. The 'contentful' uses of *do* – e.g. *She did him eat and drink* 'She caused him to eat and drink' – exhibited the same infinitive selection properties as the modals, as discussed above, producing the construction NP-*do-*

'bare' infinitive. And, indeed, periphrastic *do* also begins to appear around the end of the 15th-century, coinciding with the beginning of the reanalysis of *to* and the modals. Our analysis captures this connection as *X°MVT is ranked below V+INFL and is thus also ranked below FILL, which penalizes the appearance of semantically empty elements. Thus, in Middle English, whilst this constraint maintains its high-ranking, *do*-support is non-optimal, as shown in Tableau 12.

Tableau 12. do *in ME*

	OBHD	HMC	PRBD	FILL	V+INFL	*X°MVT
a.[TP NP **do** [MP **t** [VP V				*!		*
☞b. [TP NP **V** [MP **t** [VP **t**						**

However, development of periphrastic *do* does not entirely parallel that of the either of the latter elements. That is to say that the loss of iterating modals precedes the categorical establishment of the SPDE patterning of *do* by at least a century and a half. This divergence of the development of *do* and the other unbound inflections is unsurprising, as the use of *do*, violates not V+INFL but rather FILL, being a semantically empty element. The connection between the modals and *do* is that they both reflect the outcome of the increasing prohibition of verb-raising.

Tableau 13. do *in EME*

	OBHD	HMC	*X°MVT	FILL	V+INFL
☞a. [TP NP **do** [MP **t** [VP V			*	*	
b. [TP NP **V** [MP **t** [VP **t**			**!		

As Tableau 13 shows, when *X°MVT comes to dominate FILL we first observe the appearance of the empty periphrastic *do*. It is to an understanding of the development of *do*-support that we now turn.

3.2. A diachronic OT analysis of periphrastic do

In this section we develop a fuller analysis of *do*-support in English, accounting for the historical development of *do* across syntactic contexts. Figure 2 shows that periphrastic *do* displays different relative frequencies of use across contexts. Whilst we do not propose to account for the exact frequency of *do* at any point in time, the OT account proposed below does account for the distribution of proportion of the use of *do*. That is, it derives the fact that, for instance, negative declaratives never show a greater frequency of *do*-support than do affirmative interrogatives.

Beginning in the late 15th century, speakers begin to analyze *do* as a semantically empty element. As such, for those speakers, *do* is no longer part of the underlying semantic structure and its presence must be imposed by the syntactic constraints of

the grammar itself. Thus, language learners receive evidence apparently attesting a grammar in which semantically empty *do* is required. However, as speakers of the 'older' grammar obviously do not invariably employ *do* – which for them is 'contentful' – learners also receive conflicting evidence pointing to a grammar in which *do* is not required.

We propose that this apparently contradictory evidence leads to the situation in which language learners do not acquire a single total ranking, but rather only a partial ordering of constraints, which support multiple grammars. This is the scenario of competing grammars discussed in §2. The existence of intra-, as well as inter-, speaker variation, also discussed in §2, indicates that the increase in the use of *do* during the EME period is not simply the spread of a '*do*-grammar' through the community (though obviously part of the rise is due to the spread of new grammars through the linguistic community); part of the rise of *do*-support is due to the increasing tendency on the part of an individual speaker towards the use of *do*-constructions. Thus, we would also like to account for the increase in the use of *do* within speakers during the EME period. The rise of periphastic *do* use within speakers logically could be the result of one of (at least) two constraint rerankings. Either the constraint prohibiting *do*-support could be falling in rank or the constraint prohibiting verb-raising could be rising in rank. As discussed in §2, Kroch (1989b:223-225) provides empirical arguments that point to the latter possibility. Therefore the change that brings about *do*-support is, in our terms, the rise in rank of *X°MVT, which prohibits verb-movement. As *X°MVT comes to dominate FILL, *do*-support is chosen over verb-raising. However, if the change simply involved these two constraints we should not expect to see a difference in frequencies between contexts. What Figure 2 shows us is that there is something worse about verb-raising in interrogative contexts than in declarative contexts. The obvious difference between these two contexts is that in the former the verb need not only raise to T° but further to C° – see Tableaux 3 and 4 above.

To capture this difference we employ a formal device introduced in Smolensky (1997): *constraint conjunction*, in which the simultaneous violation of two constraints may produce a higher-ranked violation. In this instance *X°MVT is conjoined with itself, resulting in a *universal subhierarchy* (cf. Legendre et al. 1998), reflecting that multiple violations of a certain constraint are equivalent to a violation of a higher ranked constraint. In addition to the constraint *X°MVT, we also propose *X°MVT^2 and *X°MVT^3. These constraints stand in a *universally*-fixed ranking *X°MVT^3 » *X°MVT^2 » *X°MVT, and it is this universal subhierarchy which is reranked in EME, as we will see. In 3.2.1 we provide an account of the rise of *do*-support across negative and affirmative declarative and interrogative contexts up to the close of the 16th century – that is, an account of *do*-support during the period in which Kroch's Constant Rate Hypothesis mathematically holds (cf. Kroch 1989a,b).

An interesting secondary change occurs in the late 16th century – from Ellegård's data apparently circa 1575 – which causes a brief falling-off of *do*-support in the negative contexts and marks the beginning of the permanent loss of *do* from non-

emphatic affirmative declaratives. §3.2.2 provides an account of the grammatical changes underlying these effects.

3.2.1. *Periphrastic* do *in early Early Modern English*

Language learners in the EME period are presented with sentences both with and without *do* – the two types lacking any apparent semantic differentiation. Thus they acquire a grammar that allows for both types of constructions. For declaratives, this necessitates positing both the rankings *X°Mvt^2 » Fill,PrBd – generating clauses with *do*, and Fill,PrBd » *X°Mvt^2, producing clauses without *do*. Similarly, interrogatives require both the rankings *X°Mvt^3 » Fill,PrBd and Fill,PrBd » *X°Mvt^3. Thus speakers of EME maintain multiple constraint rankings, some with the *X°Mvt constraints highly ranked, some with those constraints lower ranked. As indicated by the absence of a ranking relationship between them, Fill and PrBd also alternate in rank during this period. The absence of a fixed ranking between Fill and PrBd may be due to the want of evidence attesting their relative ranking. This want is due to the fact that in earlier English neither *do*-support nor affix-lowering are grammatical options, effectively attesting only that neither is dominated by the *X°Mvt constraints, without providing data concerning their ranking with respect to each other.

The variability of ranking of Fill with respect to PrBd, in addition to the variability of the ranking of *X°Mvt constraints, is attested by the appearance of *three* clause-types in the affirmative declarative context, as discussed below: *John speaks never* [verb-raising], *John does never speak* [*do*-support], *John never speaks* [affix-lowering], violating *X°Mvt^2, Fill and PrBd, respectively.

In this section, utilizing the set of possible constraint rankings shown in Table 2, which is the full range of possible rankings generated by the variation discussed above,[13] we account for the occurrence and relative frequency of the three clause-types (*do*-support, affix-lowering, verb-raising) across contexts during the EME period, assuming the frequency of each of the rankings to be non-zero.

We lack the data to assign exact percentages to the frequency of the use of any particular rankings – recall that Ellegård's study simply provides a gross frequency of *do* use, collapsing over inter- and intra-speaker variation. Because of the inter-speaker variation, we cannot give a single definitive grammar for any one point in time, as individual speakers show idiosyncratic difference in their employment of different grammars. However, in general, we can say that, summing over the entire community, the rankings in which *X°Mvt^3, *X°Mvt^2 are relatively low ranked, with respect to PrBd and Fill, are most prevalent in the earliest part of the EME period. High ranking of *X°Mvt^3, *X°Mvt^2 becomes increasingly predominant as the period progresses, being responsible for the increased frequency of *do*-support overall. The ranking Fill » PrBd becomes more common close to 1575, and after that point becomes the predominant ranking, as discussed in §3.2.2.

Table 2. Early EME rankings

A. $*X°M\textsc{vt}^3$ » $*X°M\textsc{vt}^2$ » P\textsc{r}B\textsc{d} » F\textsc{ill}

B. P\textsc{r}B\textsc{d} » $*X°M\textsc{vt}^3$ » $*X°M\textsc{vt}^2$ » F\textsc{ill}

C. P\textsc{r}B\textsc{d} » F\textsc{ill} » $*X°M\textsc{vt}^3$ » $*X°M\textsc{vt}^2$

D. $*X°M\textsc{vt}^3$ » P\textsc{r}B\textsc{d} » $*X°M\textsc{vt}^2$ » F\textsc{ill}

E. P\textsc{r}B\textsc{d} » $*X°M\textsc{vt}^3$ » F\textsc{ill} » $*X°M\textsc{vt}^2$

F. $*X°M\textsc{vt}^3$ » P\textsc{r}B\textsc{d} » F\textsc{ill} » $*X°M\textsc{vt}^2$

G. $*X°M\textsc{vt}^3$ » $*X°M\textsc{vt}^2$ » F\textsc{ill} » P\textsc{r}B\textsc{d}

H. F\textsc{ill} » $*X°M\textsc{vt}^3$ » $*X°M\textsc{vt}^2$ » P\textsc{r}B\textsc{d}

I. F\textsc{ill} » P\textsc{r}B\textsc{d} » $*X°M\textsc{vt}^3$ » $*X°M\textsc{vt}^2$

J. $*X°M\textsc{vt}^3$ » F\textsc{ill} » $*X°M\textsc{vt}^2$ » P\textsc{r}B\textsc{d}

K. F\textsc{ill} » $*X°M\textsc{vt}^3$ » P\textsc{r}B\textsc{d} » $*X°M\textsc{vt}^2$

L. $*X°M\textsc{vt}^3$ » F\textsc{ill} » P\textsc{r}B\textsc{d} » $*X°M\textsc{vt}^2$

We shall consider four syntactic contexts in turn: affirmative declaratives, affirmative interrogatives, negative declaratives and negative interrogatives; showing that Table 2 contains the crucial ranking permutations required in order to generate the clause-types observed in the first half of the EME period, also deriving, to a certain degree of accuracy, the relative frequency of occurrence of the three options (*do*-support, affix-lower, verb-raising) across contexts.

Examples (18)[14] are representative of the three clause-types in the affirmative declarative context:

(18)　　Affirmative declaratives

　　　a.　Therefore it is meet that <u>noble minds keep ever</u> with their likes….
　　　　　(*Julius Caesar* I.ii.311-312) [verb-raising]

　　　b.　<u>The King doth keep</u> his revels here to-night;
　　　　　(*A Midsummer's Night Dream* II.i.18) [*do*-support]

　　　c.　… <u>the master I speak of ever keeps</u> a good fire.
　　　　　(*All's Well That Ends Well* IV.v.47-49)[affix-lowering]

Tableaux 14-16 display the candidate evaluations on the range of constraint rankings possible during the early part of the EME period. The clause type generated (i.e. a, b, c from (18) above) is indicated above each tableau, as are the rankings (from (19) above) that produce that clause type. The ranking(s) that the tableau itself represents are **bolded**; non-bolded rankings generated identical outcomes. This method minimizes the number of explicit tableaux required.

Note that in the following sections no single tableau represents a 'stage' of English. As we have discussed, speakers of English in the EME period show variance between a number of 'competing grammars'. Any tableaux with vertical broken lines represent a set of possible rankings – that is, a set of possible grammars. On a tableau in which constraint A and constraint B are separated by a broken rather than a solid line, this does not indicate that they are 'equally ranked', but instead that on some evaluations A » B and on others B » A. Tableaux without any broken lines represent an individual ranking – which itself may be a member of

a set of possible rankings at some 'stage' of English. That is, though on a given tableau A » B, at some stage that ranking may also alternate with B » A.

Tableau 14. EME Affirmative declaratives [type a]
*[rankings (from (17) above): **C**, E, F, **I**, K, L]*

	OB HD	HMC	PR BD	FILL	*X° MVT³	*X° MVT²	*X° MVT
☞ a. [TP NP **keep** [MP **t** [ADVP ever][VP **t**						*	**
b. [TP NP **do** [MP **t** [ADVP ever][VP keep				*!			*
c. [TP NP *t* [MP *t* [ADVP ever][VP *keep*			*!				
d. [TP NP e̱ [MP e̱ [ADVP ever][VP keep	*!*						

Tableau 14 shows an evaluation in which verb-raising (candidate 14a) is optimal. In Tableau 14, and in all subsequent tableaux, candidate d. is non-optimal due to its violation of the undominated constraint OBHD, as it leaves T° and Mood° unfilled. Candidate 14b crucially violates FILL through the insertion of *do*. Candidate c., lowering affixes from T° and Mood° onto the verb in V°, violates PRBD, which dominates *X°MVT², and thus c. is ruled non-optimal.

*Tableau 15. EME affirmative declaratives [type b] [rankings: **A**, B, D]*

	OB HD	*X° MVT³	*X° MVT²	HMC	PR BD	FILL	*X° MVT
a. [TP NP **keep** [MP **t** [ADVP ever][VP **t**			*!				**
☞ b. [TP NP **do** [MP **t** [ADVP ever][VP keep						*	*
c. [TP NP *t* [MP *t* [ADVP ever][VP *keep*					*!		
d. [TP NP e̱ [MP e̱ [ADVP ever][VP keep	*!*						

When *X°MVT² » PRBD » FILL, as in Tableau 15, *do*-support (b), rather than verb-raising or affix-lowering is optimal.

*Tableau 16. EME affirmative declaratives [type c] [rankings: **G**, H, J]*

	OB HD	*X° MVT³	*X° MVT²	HMC	FILL	PR BD	*X° MVT
a. [TP NP **keep** [MP **t** [ADVP ever][VP **t**			*!				**
b. [TP NP **do** [MP **t** [ADVP ever][VP keep					*!		*
☞ c. [TP NP *t* [MP *t* [ADVP ever][VP *keep*						*	
d. [TP NP e̱ [MP e̱ [ADVP ever][VP keep	*!*						

If the ranking of FILL, PRBD is reversed from that in Tableau 15, affix-lowering (c) is optimal, as Tableau 16 shows.

Affirmative interrogatives exhibit fewer options than affirmative declaratives, as shown in (19), we only find two clause-types: verb-raising or *do*-support:[15]

(19)   Affirmative interrogatives

   a. <u>What means our cousin</u> that he stares and looks so wildly?
    (*Richard II* V.iii.24) [verb-raising]

   b. <u>What do you mean</u> to dote thus on such luggage?
    (*Tempest* IV.i.229-230) [*do*-support]

In Tableaux 17 and 18, affix-lowering (candidate (c)) is non-optimal, for C° bears no affixes or features itself, and thus no affixes can be lowered from C°, as discussed in §1. Therefore, if affix-lowering is applied, C° is left empty, violating OBHD. Otherwise the affirmative interrogatives are very similar to affirmative declaratives, excepting that the crucial ranking pertains to FILL and $*X°MVT^3$, rather than $*X°MVT^2$ as in the declaratives.

*Tableau 17. EME affirmative interrogative [type a] [rankings: **C**, H, **I**, K]*

	OB HD	HMC	PR BD	FILL	$*X°MVT^3$	$*X°MVT^2$	$*X°MVT$
☞ a. [CP What **means** [TP NP **t** [MP **t** [VP **t**					*	*	***
b. [CP What **does** [TP NP **t** [MP **t** [VP mean				*!		*	**
c. [CP What <u>e</u> [TP NP *t* [MP *t* [VP *means*	*!		*				
d. [CP What <u>e</u> [TP NP <u>e</u> [MP <u>e</u> [VP *means*	*!*						

*Tableau 18. EME affirmative interrogative [Type b] [Rankings: A, B, D, E, **F**, G, J, **L**]*

	OB HD	$*X°MVT^3$	HMC	PR BD	FILL	$*X°MVT^2$	$*X°MVT$
☞ a. [CP What **means** [TP NP **t** [MP **t** [VP **t**		*!				*	***
b. [CP What **does** [TP NP **t** [MP **t** [VP mean					*	*	**
c. [CP What <u>e</u> [TP NP *t* [MP *t* [VP *means*	*!			*			
d. [CP What <u>e</u> [TP NP <u>e</u> [MP <u>e</u> [VP means	*!*						

Negative declaratives in the first half of the EME period are like affirmative interrogatives in displaying only two clause-types, as shown in (20):

(20)   Negative declaratives

   a. What music will be in him when Hector has knockd out his brains <u>I know not</u>; but, I am sure, none; unless the fiddler Apollo get his sinews to make catlings on. (*Troilus and Cressida* III.iii.291-294) [verb-raising]

   b. <u>I do not know</u> that Englishman alive with whom…
    (*Richard III* II.i.70-72) [*do*-support]

As discussed in §1, there is evidence for at least two different positions for sentential negation in English. The higher negation, NegP1, intervenes between TenseP and MoodP, and the lower, NegP2, between MoodP and VP. When a clause is formed with the higher negation, with *not* in the head of NegP1, as in Tableaux 19 and 20, the outcome is essentially identical to that of the affirmative declaratives as shown in Tableaux 14 and 15. The important difference is the lack of an affix-lowering type parallel to that shown in Tableau 16. The reason behind this is the double violation of the HMC by affix-lowering, which rules out candidate c. in both Tableau 19 and Tableau 20 due to the high rank of the HMC.

Tableau 19. EME negative declarative with higher negation [type a]
*[rankings: C, E, **F**, H, I, J, K, **L**]*

	OB HD	*X° MVT³	HMC	PR BD	FILL	*X° MVT²	*X° MVT
☞ a. [TP NP **know** [NEGP1 not [MP **t** [VP **t**			*			*	**
b. [TP NP **do** [NEGP1 not [MP **t** [VP know			*		*!		*
c. [TP NP *t* [NEGP1 not [MP *t* [VP *know*			**!	*			
d. [TP NP e [NEGP1 not [MP e [VP know	*!*						

In Tableau 19, candidates (a), (b) and (c) all violate the HMC. However, candidate 19c violates the HMC twice, as mentioned, because of the raising required at LF, which rules it non-optimal. Both a. and b. violate the HMC once – as both the main verb in 19a and *do* in 19b cross over NegP1 on the way to T°. These violations cancel out and the crucial ranking remains that of *X°MVT² and FILL, as in the affirmative declaratives. Therefore, when FILL » *X°MVT², as in Tableau 18, verb-raising is optimal; when *X°MVT² » FILL, as in Tableau 20 below, *do*-support is optimal.

*Tableau 20. EME Negative Declarative with higher negation [type b] [rankings: **A**, B, D, **G**]*

	OB HD	*X° MVT³	*X° MVT²	HMC	PR BD	FILL	*X° MVT
a. [TP NP **know** [NEGP1 not [MP **t** [VP **t**			*!	*			**
☞ b. [TP NP **do** [NEGP1 not [MP **t** [VP know				*		*	*
c. [TP NP *t* [NEGP1 not [MP *t* [VP *know*				**!	*		
d. [TP NP e [NEGP1 not [MP e [VP know	*!*						

However, if *not* resides in NegP2, between MoodP and VP, *do*-support is always optimal, regardless of the relative ranking of *X°MVT² with respect to FILL.

Tableau 21. EME negative declarative with lower negation [type b]
*[rankings: C, E, **F**, H, I, J, K, **L**]*

	OB HD	*X°MVT^3	HMC	PR BD	FILL	*X°MVT^2	*X°MVT
a. [TP NP **know** [MP t [NEGP2 not [VP t			*!			*	**
☞b. [TP NP **do** [MP t [NEGP2 not [VP know					*		*
c. [TP NP t [[MP t [NEGP2 not [VP *know*			*!*	*			
d. [TP NP e [MP e [NEGP2 not [VP know	*!*						

In Tableau 21, if the HMC were not active, as in the affirmative declarative context with the identical constraint ranking as in Tableau 14, verb-raising would be optimal. However, as *do* originates in Mood°, and, as such, does not cross over NegP2 on its way to T°, the use of *do* avoids a violation of the HMC when *not* occupies the lower negation position. In contrast, the verb-raising candidate, 21a, violates the HMC as the main verb crosses over NegP2 on the step between V° and Mood°. Thus, the ranking of *X°MVT^2 with respect to FILL is irrelevant when the negative is formed with NegP2, as Tableaux 21 and 22 demonstrate. Hence, rather, the crucial ranking is that of HMC and FILL, because of the asymmetry between the use of *do* in negatives with NegP1 as opposed to NegP2.

*Tableau 22. EME negative declarative with lower negation [type b] [rankings: A, B, D, **G**]*

	OB HD	*X°MVT^3	*X°MVT^2	HMC	PR BD	FILL	*X°MVT
a. [TP NP **know** [MP t [NEGP2 not [VP t			*!	*			**
☞b. [TP NP **do** [MP t [NEGP2 not [VP know						*	*
c. [TP NP t [[MP t [NEGP2 not [VP *know*				*!*	*		
d. [TP NP e [MP e [NEGP2 not [VP know	*!*						

Negative interrogatives throughout the EME period utilize only either verb-raising or *do*-support, as shown in the examples in (21):

(21)   Negative interrogatives in EME

   a.   <u>Know'st thou not</u> that I have fined these bones of mine for ransom?
        (*Henry V* IV.vii.67-68) [verb-raising]

   b.   <u>Do not you know</u> my lady's foot by th' squier…
        (*Love's Labour's Lost* V.ii.474-475) [*do*-support]

The evaluations in Tableaux 23-26 are identical to those in Tableaux 19-22, except that the crucial ranking is between *X°MVT^3, rather than *X°MVT^2, and FILL. The

same situation concerning negatives with NegP2 holds in interrogatives, as shown in Tableaux 25 and 26.

*Tableau 23. EME negative interrogative with higher negation [type a] [rankings: **C**, H, **I**, K]*

	OB HD	HMC	PR BD	FILL	*X° MVT³	*X° MVT²	*X° MVT
☞a. [CP Op **know** [TP NP **t** [NEGP1 not [MP **t** [VP **t**		*			*	*	***
b. [CP Op **do** [TP NP **t** [NEGP1 not [MP **t** [VP know		*		*!		*	**
c. [CP Op e [TP NP t [NEGP1 not [MP t [VP *know*	*!	**					
d. [CP Op e [TP NP e [NEGP1 not [MP e [VP know	*!*						

Tableau 24. EME negative interrogative with higher negation [type b]
*[rankings: A, B, D, E, **F**, G, J, **L**]*

	OB HD	*X° MVT³	HMC	PR BD	FILL	*X° MVT²	*X° MVT
a. [CP Op **know** [TP NP **t** [NEGP1 not [MP **t** [VP **t**		*!	*			*	***
☞b. [CP Op **do** [TP NP **t** [NEGP1 not [MP **t** [VP know			*		*	*	**
c. [CP Op e [TP NP t [NEGP1 not [MP t [VP *know*	*!		**				
d. [CP Op e [TP NP e [NEGP1 not [MP e [VP know	*!*						

*Tableau 25. EME negative interrogative with lower negation [type b] [rankings: **C**, H, **I**, K]*

	OB HD	HMC	PR BD	FILL	*X° MVT³	*X° MVT²	*X° MVT
a. [CP Op **know** [TP NP **t** [MP **t** [NEGP2 not [VP **t**		*!			*	*	***
☞b. [CP Op **do** [TP NP **t** [MP **t** [NEGP2 not [VP know				*		*	**
c. [CP Op e [TP NP t [MP t [NEGP2 not [VP ***know***	*!	*	*				
d. [CP Op e [TP NP e [MP e [NEGP2 not [VP know	*!*						

Tableau 26. EME negative interrogative with lower negation [type b]
*[rankings: A, B, D, E, **F**, G, J, **L**]*

	OB HD	*X° MVT³	HMC	PR BD	FILL	*X° MVT²	*X° MVT
a. [CP Op **know** [TP NP **t** [MP **t** [NEGP2 not [VP **t**		*!	*			*	***
☞b. [CP Op **do** [TP NP **t** [MP **t** [NEGP2 not [VP **know**					*	*	**
c. [CP Op e [TP NP t [MP t [NEGP2 not [VP *know*	*!		*				
d. [CP Op e [TP NP e [MP e [NEGP2 not [VP know	*!*						

We have demonstrated how the patterning of *do* use in the early part of the EME period can be accounted for if speakers vary in their employment of the rankings shown in (19). The data suggest that, during the 16th century, the higher rankings of *X°MVT become increasingly dominant, producing the 'S'-like trajectory of

increasing *do* use. In the next subsection we examine the late 16th-century deviation from the 'S'-curve development of *do*-support and the loss of *do* from affirmative declaratives.

3.2.2. Periphrastic do *in late Early Modern English*

The close of the 16th century reveals a disturbance in the basic 'S'-curve pattern of linguistic renewal. During the latter part of the 16th century we observe a falling-off of *do*-support in several contexts (see Figure 2) as discussed previously. We can account for this deviation via the reranking of FILL, as this section demonstrates. This second change in the grammar occurs before the obsolescence of the rankings in which *X°MVT is low ranked and thus disrupts the 'S'-curve pattern significantly.

Additionally, the reranking of FILL accounts for another development of the period. Though the relation is not superficially obvious, we shall show that this reranking also provides an explanation of the first significant appearance of a third type of clause in the negative declarative context: the type *John not reads*. As Visser remarks, "[b]efore 1500 [the type *John not reads*] is only sporadically met with, but after 1500 its currency increases and it becomes pretty common in Shakespeare's time. After c1700 a decline sets in....[but] [t]he frequency of the use of this type in the 16th and 17th century seems to have escaped the attention of grammarians" (1963-73:§1440). This observation is confirmed by Kroch (1989b), who, however, argues against Visser's proposal that in such constructions *not* occupies a 'pre-INFL' position (Kroch's terminology): "Visser gave 59 examples of *not* before a tensed verb and in all of them, the tensed verb is a main verb. There are no cases with modals or aspectual auxiliaries and not even any with the main verbs *be* or *have*. Since the verb *to be* is by so much the most frequent verb in texts, the absence of it and other auxiliaries from Visser's list of examples is telling" (235). Kroch instead takes the type *John not reads* to be evidence of affix-lowering, as do we (see Van Gelderen 2000 for a different analysis). If the *John not reads* type were constructed with a NegP dominating TenseP (i.e. equivalent to Kroch's pre-INFL negation), then we should also expect to find clauses of the sort *John not is here*, which are conspicuously absent. Thus, we repeat the negative declaratives examples in (20), alongside of the *John not reads* type in c.

(20)   **Negative declaratives**

   a. What music will be in him when Hector has knock'd out his brains <u>I know not</u>; but, I am sure, none; unless the fiddler Apollo get his sinews to make catlings on. (*Troilus and Cressida* III.iii.291-294) [verb-raising]

   b. <u>I do not know</u> that Englishman alive with whom…
      (*Richard III* II.i.70-72) [*do*-support]

   c. **Whe'er thou be'st he or no, or some enchanted trifle to abuse me, as late I have been, <u>I not know.</u>**
      (*Tempest* **V.i.113-115) [affix-lowering]**

In affirmative declaratives, affix-lowering is optimal only when *X°MVT², FILL »
PRBD – however we saw above in Tableaux 19-22 that in negative declaratives this
ranking alone is insufficient, as the violations of the HMC rule affix-lowering non-
optimal. In order for the affix-lowering type to surface in the negative declarative
context, FILL must dominate the HMC, as in Tableau 27:

Tableau 27. Late EME neg. declarative with lower negation [type c]
*[rankings: **M**, N, P – see (25) below]*

	OB HD	*X° MVT³	*X° MVT²	FILL	HMC	PR BD	*X° MVT
a. [TP NP **know** [MP **t** [NEGP2 not [VP **t**			*!		*		**
b. [TP NP **do** [MP **t** [NEGP2 not [VP know				*!			*
☞ c. [TP NP *t* [[MP *t* [NEGP2 not [VP *know*					**	*	
d. [TP NP e [MP e [NEGP2 not [VP know	*!*						

Despite the dual violation of the HMC by candidate c., as the HMC is dominated by
both *X°MVT² and FILL, affix-lowering is the optimal type in Tableau 27. When
*X°MVT² » PRBD and FILL » PRBD, HMC, affix-lowering is optimal, regardless of
whether the higher or lower negation is chosen.

We remark at this point, that although the *John not reads* type sporadically
appears in early EME, it is clear it cannot have made up a very large percentage of
negative declarative forms prior to 1575 or else Kroch (1989b) would not have been
able to mathematically show its slope in Figure 2 – representing the frequency of the
use of *do* over time – to be essentially identical to the slope of the curves in
interrogatives. Roberts argues that the small number of apparent early instances of
this type can be analyzed as Stylistic-Fronting of *not* rather than affix-lowering (see
Roberts 1993:303-305 for details). Thus we can link the drop in frequency of *do* use
in negative contexts to the appearance of *John not reads* – the latter seems only to
occur with any frequency around and after the time at which we observe the falling-
off of *do*-support in negatives. If *John not reads* negative declaratives had made up
any significant percentage prior to 1575 then we should expect that the negative
declarative context would need to be evaluated like the affirmative declarative
context, that is, it would only be the combined frequency of the use of *do* and affix-
lowering in negative declaratives that would be equivalent to the frequency of *do*
support in the interrogative contexts.

The introduction of a set of rankings in which FILL » HMC at the close of the
16th-century, deriving the appearance of affix-lowering in negative declaratives, also
accounts for the otherwise puzzling falling-off of *do*-support in negative contexts in
general. First, let us give the full set of relevant active rankings in the late EME
period, showing the ranking of the HMC, as we shall employ in the same manner as
in the above subsection to explain the pattern of alternations in the latter half of the
EME period. A.-L. are identical to A.-L. in (25) above, but here we show the
ranking of the HMC with respect to the other constraints. Rankings M.-R. are the
'new' rankings, i.e. those in which FILL » HMC:

Table 3. Late EME rankings

A. $*X°Mvt^3$ » $*X°Mvt^2$ » PRBD » FILL　　　**G.** $*X°Mvt^3$ » $*X°Mvt^2$ » FILL » PRBD

B. PRBD » $*X°Mvt^3$ » $*X°Mvt^2$ » FILL　　　**H.** FILL » $*X°Mvt^3$ » $*X°Mvt^2$ » PRBD

C. PRBD » FILL » $*X°Mvt^3$ » $*X°Mvt^2$　　　**I.** FILL » PRBD » $*X°Mvt^3$ » $*X°Mvt^2$

D. $*X°Mvt^3$ » PRBD » $*X°Mvt^2$ » FILL　　　**J.** $*X°Mvt^3$ » FILL » $*X°Mvt^2$ » PRBD

E. PRBD » $*X°Mvt^3$ » FILL » $*X°Mvt^2$　　　**K.** FILL » $*X°Mvt^3$ » PRBD » $*X°Mvt^2$

F. $*X°Mvt^3$ » PRBD » FILL » $*X°Mvt^2$　　　**L.** $*X°Mvt^3$ » FILL » PRBD » $*X°Mvt^2$

M. $*X°Mvt^3$ » $*X°Mvt^2$ » FILL » HMC » PRBD

N. FILL » HMC » $*X°Mvt^3$ » $*X°Mvt^2$ » PRBD

O. FILL » HMC » PRBD » $*X°Mvt^3$ » $*X°Mvt^2$

P. $*X°Mvt^3$ » FILL » HMC » $*X°Mvt^2$ » PRBD

Q. FILL » HMC » $*X°Mvt^3$ » PRBD » $*X°Mvt^2$

R. $*X°Mvt^3$ » FILL » HMC » PRBD »$*X°Mvt^2$

Thus affix-lowering in negative declaratives is optimal on rankings M, N, P regardless of whether *not* occupies the head of NegP1 or NegP2.

We observe that when FILL » HMC, the asymmetry between NegP1 and NegP2, with respect to *do* use, disappears in both negative declaratives and negative interrogatives, as shown in Tableaux 28, and 29.

*Tableau 28. Late EME negative declarative with lower negation [type a] [rankings: O, Q, **R**]*

	OB HD	$*X°Mvt^3$	FILL	HMC	PR BD	$*X°Mvt^2$	$*X°Mvt$
☞a. [$_{TP}$ NP **know** [$_{MP}$ **t** [$_{NEGP2}$ not [$_{VP}$ **t**				*		*	**
b. [$_{TP}$ NP **do** [$_{MP}$ **t** [$_{NEGP2}$ not [$_{VP}$ know			*!				*
c. [$_{TP}$ NP *t* [[$_{MP}$ *t* [$_{NEGP2}$ not [$_{VP}$ *know*				**!	*		
d. [$_{TP}$ NP <u>e</u> [$_{MP}$ <u>e</u> [$_{NEGP2}$ not [$_{VP}$ know	*!*						

Since FILL » HMC, the decision between verb-raising, affix-lowering and *do*-support is made purely through the ranking of $X°Mvt^2$, the HMC and FILL. Thus, in Tableau 28, as FILL » $*X°Mvt^2$, verb-raising is optimal, despite the fact that *do*-support avoids a violation of the HMC.

Tableau 29. Late EME negative interrogative with lower negation [type a]
*[rankings: N, **O**, Q]*

	OB HD	FILL	HMC	PR BD	*X° MVT³	*X° MVT²	*X° MVT
☞a. [CP Op **know** [TP NP **t** [MP **t** [NEGP2 not [VP **t**			*		*	*	***
b. [CP Op **do** [TP NP **t** [MP **t** [NEGP2 not [VP know		*!				*	**
c. [CP Op e̲ [TP NP *t* [MP *t* [NEGP2 not [VP *know*	*!		**	*			
d. [CP Op e̲ [TP NP e̲ [MP e̲ [NEGP2 not [VP know	*!*						

Likewise, in Tableau 30, the crucial ranking is between FILL and *X°MVT². Note that the dominance of the HMC by FILL does not change the ungrammaticality of affix-lowering in negative interrogatives, as it is ruled out by OBHD, as in affirmative interrogatives.

Of course, the mere existence of the rankings in M.-R. does not directly derive the exact pattern of the drop in *do*-support in negative contexts, nor does it account for the falling-off of *do*-support in affirmative declaratives. However, the rise in rank of FILL over the HMC seems to be part of a general promotion of FILL – which includes a promotion of FILL over PRBD. After the end of the 16th century, in other words, rankings in which PRBD » FILL become rarer, until, in SPDE, FILL always dominates PRBD, ruling out *do*-support in non-emphatic affirmative declaratives. It is clear that the rankings in which FILL » HMC do not survive into the SPDE period, as attested by the lack of the clause-type *John not reads*.

3.3. Partial ordering and the distribution of do in EME

In the above sections we have shown the permutations of the constraint ranking necessary to derive the different types of constructions in our set of considered contexts in EME. Our analysis within OT allows us to characterize the basic pattern of establishment of periphrastic *do* by means of two constraint rerankings: The first is the promotion of the *X°MVT constraints, which increases the penalty for verb-raising, accounting for both the appearance of the modal class and infinitival *to*, as well as the rise of *do*-support. The second is the promotion of FILL over PRBD (and over the HMC for a time, resulting in a brief occurrence of *John not reads* clauses), which results in the ultimate loss of *do*-support from unemphatic affirmative declaratives. The general 'S'-curve pattern of increasing frequency in the use of *do*-support in the 16th and 17th centuries, however, is due to the promotion of the *X°MVT constraints, which provides us with a coherent picture of the course of change.

Our analysis also accounts for the relative proportions of *do*-support across contexts over time. First consider tables Tables 4-7, which summarize the relevant permutations of the constraint ranking during the EME for the contexts considered. The column headings show the relative positioning of the *X°MVT family of constraints w.r.t. FILL/PRBD. The row headings represent the combined alternations of the relative ranking of FILL and PRBD and HMC and FILL. The cell at the

intersection of a given column and row shows the alternate (verb-raising, *do*-support, affix-lowering) type chosen on the grammar whose variables are given by the column and row headings. In some cases, a cell may be split when the use of higher or lower negation affects the choice of alternate. The lower section of each table tabulates the total number of tableaux on which a particular alternate is chosen (out of twelve possible permutations):

V=verb-raising, D=*do*-support, A=affix-lowering

Table 4. Negative interrogatives

	FILL/PRBD » $*X°MVT^3$	$*X°MVT^3$ » FILL/PRBD » $*X°MVT^2$	$*X°MVT^2$ » FILL/PRBD » $*X°MVT$	
PRBD » FILL, HMC » FILL	**V [NegP1]** / **D [NegP2]**	**D**	**D**	
FILL » PRBD, HMC » FILL	**V [NegP1]** / **D [NegP2]**	**D**	**D**	
PRBD » FILL, FILL » HMC	**V**	**D**	**D**	
FILL » PRBD, FILL » HMC	**V**	**D**	**D**	
				Total
verb-raising	3/4	0/4	0/4	3/12
do-support	1/4	4/4	4/4	9/12
aff.-lower	0/4	0/4	0/4	0/12

Table 5. Affirmative interrogatives

	FILL/PRBD » $*X°MVT^3$	$*X°MVT^3$ » FILL/PRBD » $*X°MVT^2$	$*X°MVT^2$ » FILL/PRBD » $*X°MVT$	
PRBD » FILL, HMC » FILL	**V**	**D**	**D**	
FILL » PRBD, HMC » FILL	**V**	**D**	**D**	
PRBD » FILL, FILL » HMC	**V**	**D**	**D**	
FILL » PRBD, FILL » HMC	**V**	**D**	**D**	
				Total
verb-raising	4/4	0/4	0/4	4/12
do-support	0/4	4/4	4/4	8/12
aff.-lower	0/4	0/4	0/4	0/12

Table 6. Negative declaratives

	FILL/PRBD » *X°MVT^3	*X°MVT^3 » FILL/PRBD » *X°MVT^2	*X°MVT^2 » FILL/PRBD » *X°MVT	
PRBD » FILL, HMC » FILL	**V [NegP1]** **D [NegP2]**	**V [NegP1]** **D [NegP2]**	**D**	
FILL » PRBD, HMC » FILL	**V [NegP1]** **D [NegP2]**	**V [NegP1]** **D [NegP2]**	**D**	
PRBD » FILL, FILL » HMC	**V**	**V**	**D**	
FILL » PRBD, FILL » HMC	**V**	**V**	**A**	
				Total
verb-raising	3/4	3/4	0/4	6/12
do-support	1/4	1/4	3/4	5/12
aff.-lower	0/4	0/4	1/4	1/12

Table 7. Affirmative declaratives

	FILL/PRBD » *X°MVT^3	*X°MVT^3 » FILL/PRBD » *X°MVT^2	*X°MVT^2 » FILL/PRBD » *X°MVT	
PRBD » FILL, HMC » FILL	**V**	**V**	**D**	
FILL » PRBD, HMC » FILL	**V**	**V**	**A**	
PRBD » FILL, FILL » HMC	**V**	**V**	**D**	
FILL » PRBD, FILL » HMC	**V**	**V**	**A**	
				Total
verb-raising	4/4	4/4	0/4	8/12
do-support	0/4	0/4	2/4	2/12
aff.-lower	0/4	0/4	2/4	2/12

Table 8 conveniently compares, for each context, the relative fraction of tableaux on which a particular type is chosen:

Table 8. Contexts compared

	Verb-Raising	*do*-Support	Affix-Lowering
Neg. Int.	3/12	9/12	0/12
Aff. Int.	4/12	8/12	0/12
Neg. Decl.	6/12	5/12	1/12
Aff. Decl.	8/12	2/12	2/12

These fractions do not directly map onto frequencies of types, nor do we claim that all rankings appear with equal frequency, as discussed above. The establishment of *do*-support is partially regulated by social factors (see Stein 1990, Nurmi 1999): a particular ranking may come to be associated with certain social factors, e.g. some rankings are used in formal registers, some in informal registers, etc. Therefore, we should not expect the fractions to directly correlate with the frequency of an alternate type at a particular time. However, we can arrive at a rough approximation of the *do* frequencies over time through the competition of rankings, in which one set of rankings gradually replaces another over time. E.g. if we assume that the overall tendency in the linguistic community in the early 1500's is to grammars with lower rankings of *X°Mᴠᴛ and that this trend reverses itself as the EME period progresses, then we derive the basic 'S'-curve pattern of replacement of non-*do* forms by *do*-forms.

Further, we can additionally derive the relative proportions of *do*-use across contexts, as crucially Tables 5-8 make clear the existence of a superset/subset relation between the alternate type chosen by a particular ranking in a particular context. The rankings, for instance, on which *do*-support will be chosen as optimal in affirmative interrogative contexts constitute a subset of the rankings on which *do*-support will be chosen as optimal in negative interrogative contexts (compare Tableaux 31 and 30) and a superset of the rankings on which *do*-support will be chosen as optimal in affirmative declaratives (compare Tableaux 31 and 33).[16] Thus, at any given point in time, if we observe *do*-support occurring in 40% of negative declaratives, we can predict with confidence that *do*-support will occur at no less than 40% in affirmative interrogative and at no greater than 40% in affirmative declaratives.

3.4. Optimality Theory vs. Principles & Parameters

The coherence of the account of *do*-support present herein is made possible by the formal properties of the Optimality-Theoretic framework, in particular the interaction of violable principles. In contrast, a framework relying on a set of inviolable principles (albeit with the ability to account for crosslinguistic difference through the use of parameters or binary choice between strong and weak features in the Minimalist Program) encounters difficulties in presenting a coherent picture of a change of the complexity of the establishment of periphrastic *do*.

To illustrate, let us consider a recent analysis of the rise of *do*-support within the P&P framework, which, like ours, recognizes the validity of the Constant Rate Effect. Han & Kroch (2000) present a diachronic account of *do*-support in terms of competing grammars, under a P&P formalism. The following discussion is not intended as a criticism of Han & Kroch's analysis *per se* – indeed we have taken advantages of many of their insights, such as the multiple positions of negation and their consequences, as well as having followed in the general theory of grammatical competition developed in Kroch (1989a,b) – but rather to point of the limitations of the P&P framework in which they cast their analysis.

Han & Kroch assume the following clause structure, similar to the one we assumed in our analysis, excepting for the addition of an Aspect projection between MoodP and NegP2:
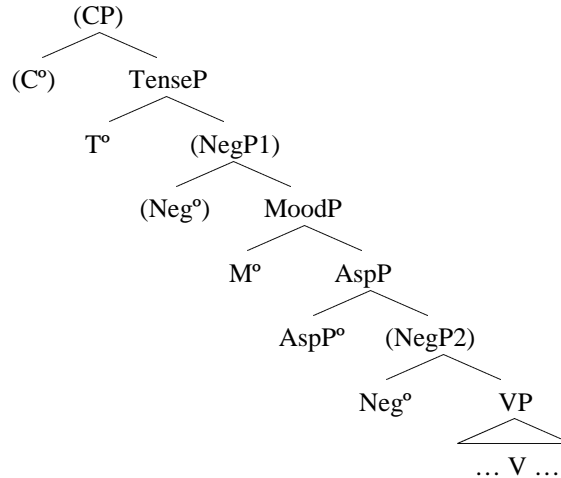
```
              (CP)
             /    \
          (C°)    TenseP
                 /      \
               T°      (NegP1)
                      /       \
                  (Neg°)     MoodP
                            /     \
                          M°      AspP
                                 /    \
                             AspP°   (NegP2)
                                     /      \
                                  Neg°      VP
                                           /  \
                                        … V …
```

Figure 6. Phrase structure (Han & Kroch 2000).

Han & Kroch propose that there are two stages in the establishment of *do*-support. The first stage they correlate to the gradual loss of Mood-to-Tense movement. In many contexts, as described below, this loss necessitates the use of *do*-support. Therefore, it is in this stage that the meaningless periphrastic *do* first appears, and competes with the conservative option of verb-raising. The second stage they align with the progressive loss of V-to-Aspect movement. The obsolescence of grammars with V-to-Aspect movement produces the SPDE patterning of *do*-support across contexts. Thus, Han & Kroch derive the loss of verb-raising in the history of English syntax, and its replacement by *do*-support, via the loss of these two formally-disjoint grammatical operations from English grammars.

From 1500-1575, Han & Kroch (2000) propose a competition between a grammar with Mood-to-Tense movement and one lacking that operation. Presumably, in Minimalist terms, this indicates something of the order of a competition between a grammar with either a strong V- or a strong Mood-feature in T° (Han & Kroch make no specific reference to movement in terms of features, so the choice between the two is uncertain), requiring checking via movement of a verb to T° before Spell-out, and one in which the V-feature (or Mood feature) in T° is weak, thus allowing verbs to check this feature after Spell-out at LF. Han & Kroch term the latter type of operation feature – rather than category – raising, in our terms, equivalent to affix-lowering. They further assume that sentential negation (i.e. *not*) blocks feature movement, but not category (head) movement. Thus, the higher negation blocks feature movement from Mood-to-Tense, and the lower from Asp-to-

Mood. One upshot of this is that when a grammar without the Mood-to-Tense operation is employed, if the higher negation is chosen, *do*-support is required as checking of the Tense feature in T° via feature-raising is blocked by the intervening negation.

Interrogatives, on this account, have a strong [+Q] operator in C that needs to be checked before Spell-out, thus always requiring an overt element to move to C° at Spell-out. Therefore, whenever the grammar lacking Mood-to-Tense movement is employed, *do*-support is also required. It is unclear, however, why *do* is ever found in affirmative declaratives where nothing stands in the way of feature raising.

Han & Kroch propose that the loss of Mood-Tense movement reaches completion circa 1575. This proposal raises a number of issues, since we do not find categorical use of *do* in interrogatives at this time. Han & Kroch address this by supposing that a third grammar exists during the early EME and extending into the later part of that period: one lacking Mood-to-Tense movement, but with direct Asp-to-C movement. Han & Kroch appeal to a P&P analysis of V2 languages (Platzack & Holmberg 1989) that employs direct V-to-C movement, skipping over intermediate heads in their proposal of 'long-distance' Asp-to-C movement.

In summary, there are three competing grammars given in (26) along with that of SPDE, producing the clause-types as in (26) [category-mvt=our verb-raising; φ-mvt (feature movement)=our affix-lowering]; *do*-support=our *do*-support]:

Table 9. Competing systems in Han & Kroch (2000)

Contexts	Grammars			
	+ Mood-to-Tense + V-to-Asp - Asp-to-C	**- Mood-to-Tense + V-to-Asp - Asp-to-C**	**- Mood-to-Tense + V-to-Asp + Asp-to-C**	**- Mood-to-Tense - V-to-Asp ± Asp-to-C[17]**
Neg. Int.	category-mvt	*do*-support	category-mvt	*do*-support
Aff. Int.	category-mvt	*do*-support	category-mvt	*do*-support
Neg. Decl. [high neg.]	category-mvt	*do*-support	*do*-support	*do*-support
Neg. Decl. [low neg.]	category-mvt	category-mvt[18]	category-mvt[17]	*do*-support
Aff. Decl.	category-mvt	φ-mvt?[19]	φ-mvt?[18]	φ-mvt
Dates	c.1500-c.1575	c.1500-? (c.1700?)	c.1500-? (c.1700?)	c.1575-Present Day

Thus, from 1500 to 1575, three grammars compete, one with Mood-to-Tense movement, and two without; one of these latter possesses a direct Asp-to-C operation lacked by the other two grammars. Around 1575, the two grammars without Mood-to-Tense movement win at the expense of the grammar with Mood-to-Tense movement – the existence of the grammar with Asp-to-C movement accounting for the continual appearance of verb-raising in interrogatives after 1575. Sentences of the type *John reads not* continue to appear after 1575 only when the lower negation is employed, on grammars still possessing V-to-Asp movement: the main verb raises to Asp, linearly preceding *not*, with subsequent feature-raising to Mood° and Tense°, unblocked by the lower negation.

From 1575 onwards, V-to-Asp movement is gradually lost, causing the loss of verb-movement in all contexts, making *do*-support obligatory in negatives, as, if the verb remains in V°, feature-raising is blocked whether the higher or lower negation is employed. And in interrogatives, *do*-support is obligatory with the loss of V-to-Asp movement due to the strong [+Q] feature in C°, as even direct Asp-to-C is insufficient to raise the verb to C° if the verb is unable to raise first to Asp°.

This account is largely consistent with the observed data (barring a few questions, such as the reason for *do*-support in affirmative declaratives on this account) – in addition offering explanation for the late appearance of *do*-support in negative imperatives, which we do not cover here (see Han & Kroch 2000, Han 2000 for details). However, our OT-based analysis compares favorably to the P&P account in several regards.

Firstly, there is the question of the 'long-distance' movement from Asp-to-C. Even setting aside the general undesirability of allowing 'head skipping', there is the question of aligning principles with movement. It is not clear what role the strength of features plays in regard to such movement, i.e. what are the principles requiring or allowing direct Asp-to-C movement in some cases and disallowing it in others?

The main aspect in which an OT analysis is superior to its P&P counterpart is the overall conceptualization and the theoretical machinery required to describe the course of a change like that of the establishment of periphrastic *do*.[20] In the analysis of Han & Kroch (2000) the progressive loss of verb-raising operations and corresponding increase in use of *do*-support involves three essentially unrelated changes: the loss of the Mood-to-Tense operation (perhaps due to the loss of a strong V- or Mood- feature in T°), the innovation and subsequent loss of a 'direct' Asp-to-C operation (with no obvious featural motivation), and the loss of the V-to-Asp operation (perhaps due to the loss of a strong V-feature in Asp°). The overall effect of these changes brings about the loss of verb-movement, but the changes are not formally or logically related. That is, on a parametric view of crosslinguistic difference, one language may have a strong Mood-feature in T° and a weak V-feature in Asp° and another may have a weak Mood-feature in T° and a strong V-feature in Asp°. In this way, the 'S'-curve pattern of linguistic renewal, which intuitively reflects the replacement of one set of structures by another, is split into a number of unrelated changes, i.e. there is an unexplained 'conspiracy'. The first half of the 'S'-curve is related to one set of changes and the second half to another.

In contrast, our OT analysis accounts for the systematic course of the 'S'-curve development of periphrastic *do* by an equally systematic development in the underlying grammars. The rise of *do*-support is correlated with the promotion of the *X°MVT constraints; the variation is accounted for by positing variance in their ranking during the EME period. The 'S'-curve style increase in the use of *do* reflects the gradual loss of the lower rankings of the *X°MVT constraints, i.e. *X°MVT^x comes to dominate FILL on a higher percentage of occasions. In addition, our analysis accounts for the relative proportions of *do*-use across contexts. As the lower rankings of the *X°MVT constraints disappear, interrogatives tend to show *do*-use more often than do declaratives, for it will be more often the case that *X°MVT^3 »

FILL (necessary for *do*-support in interrogatives) than it is the case that $*X°M\text{VT}^2$ » FILL (necessary for *do*-support in declaratives).

Furthermore, the OT analysis resolves another debate concerning the relation between verb-movement and the development of the 'modals'. Lightfoot (1979), Kroch (1989b) and Roberts (1993) conclude that V-to-I movement is lost prior to the 17th century. Tieken-Boon van Ostade (1987), Warner (1997) and Lightfoot (1999) date the loss much later, Tieken-Boon van Ostade offering evidence of verb-raising into the 18th century. Kroch (1989b) and Roberts (1993) link the development of the modals to loss of V-to-I movement – an appealing position, as the changes intuitively seem related. However, the evidence that verb-raising movement continues into the 18th century makes this an extremely questionable position, as the consensus of Lightfoot (1979 et seq.), Kroch (1989), Roberts (1985, 1993) and Warner (1993) is that change bringing about the SPDE status of the 'modals' is complete by the early 16th century. Roberts (2000), in fact, goes so far in separating the two changes as to conclude that the reanalysis of the modals is a distinct change from the loss of V-to-I movement altogether, involving sporadic reanalysis of individual modals as highly-specific functional projections (in the sense of Cinque 1999).

Whilst recognizing the existence of verb-raising in the 18th century (Lightfoot 1999, Warner 1997, Tieken-Boon van Ostade 1987), we can yet link the development of the modals to the rise of *do*-support and the parallel loss of verb-raising. The crucial constraints concerning the modals are $*X°M\text{VT}$ and V+INFL; those crucial for *do*-support are $*X°M\text{VT}$ and FILL (and PRBD for affix-lowering). The link between the development of the modals and the rise of *do* is the promotion of the $*X°M\text{VT}$ constraints. The temporal discrepancy between the completion of the reanalysis of the modals and the loss of verb-raising is due to the fact that grammars in which V+INFL » $*X°M\text{VT}$ (the 'pre-modal' grammars) become obsolete before do the grammars in which FILL » $*X°M\text{VT}^X$ (verb-raising grammars). Both changes involve an increasing price for verb-movement; they differ in the relative cost of the competing principle. With a relatively small set of constraints and merely two constraint rerankings we are able to account for progressive loss of verb-raising in favor of *do*-support and, in some contexts, affix-lowering (gradual domination of FILL, PRBD by the $*X°M\text{VT}$ constraint-family), as well as the late 16th-century deviation from the 'S'-curve pattern of *do* use and loss of *do* from affirmative declaratives (promotion of FILL). The P&P analysis of Han & Kroch (2000) requires positing a set of logically unrelated changes in order to account for a relatively consistent course of development of surface structures.

As Grimshaw (1993) stated at the outset of the Optimality-Theoretic program: "Maximally general principles will inevitably conflict. The alternative is to formulate more specific principles that are designed never to conflict, and the price is generality" (cited in Burzio 1998:111). Our OT analysis takes advantages of general principles, such as economy of movement, and demonstrates how a simple re-ordering of the relative precedence of these general principles can describe a

change such as the rise of *do*-support – as well as linking it to other developments in the history of English, such as that of the modals and infinitival *to*.

4. CONCLUSION

Descriptively, the 'S'-curve pattern of language change reflects the replacement of one linguistic form or structure by another. The notion of competing grammars, invoked by Kroch (1989a,b; 1994) and followed in our analysis, relates this course of replacement to the gradual ousting of one grammar by another. The case of *do*-support instantiates this situation but is further complicated by the deviation from the 'S'-curve pattern that occurs sometime around 1600. At that time, the use of *do*-support falls off in the negative contexts and in affirmative declaratives – never to recover in the case of the latter. This deviation appears to be the result of a second change that cuts across the pattern of the first, *do*-'favoring', change. Even aside from the latter deviation, the complexity of fluctuating pattern of *do*-use, verb-raising, and affix-lowering in the EME period raises the question of whether or not one can plot the fairly systematic descriptive course of the development of periphrastic *do* in terms of a unitary development of the underlying grammars.

Our OT analysis accounts for the patterning of clause alternates – e.g. *do*-support, verb-raising, affix-lowering – at various points during the EME period by means of merely two constraint rerankings. The first is the promotion of the *X°MVT constraints, which, by coming to dominate V+INFL, cause the appearance of unbound inflection, e.g. the 'modals' and infinitival *to*; and, as they are promoted over FILL, also trigger the use of periphrastic *do*, as well as affix-lowering, by dominating PRBD. This accounts for the roughly contemporaneous appearance of the meaningless periphrastic *do*, the 'modal' class and the use of *to* as a marker of infinitival mood. The promotion of the *X°MVT constraints accounts for the rise of *do*-support across contexts *at the expense of verb-raising*, in accordance with the theory of grammatical competition, in which one construction increases in frequency in direct proportion to the decrease in frequency of use of another (Kroch 1989a,b; 1994). The second reranking, the promotion of FILL, distorts the 'S'-curve pattern of the first change and produces the loss of *do*-support from affirmative declaratives – thus the fleeting depression of *do*-use in negative contexts, as well as the brief appearance of the clause type *John not reads*.

We have shown that, by means of *partial constraint ordering*, we are able to formally express the variation between clause-types observed in the EME period within the OT framework. Partial constraint ordering allows for a formal encoding of the overall similarity of alternating grammars, i.e. most of the dominance relations between constraints remain constant over all ranking, while representing the existence of fluctuating structures within the linguistic community. Speakers appear able to maintain multiple constraint rankings, which are associated, in some cases, with sociolinguistic factors, to wit, certain rankings align with 'high' styles and others with 'low' styles. The establishment and loss of grammars seems at least partially determined by fluctuating sociolinguistic features associated with said

grammars. We have argued that language acquisition plays the greatest role in the establishment of *new* grammars, sociolinguistic factors being relevant only once a grammar is made available and is established within the linguistic community.

Further, interaction of constraints inherent in the OT framework allows for a derivation of the relevant proportions of *do* use across contexts as different constraints are crucial in different contexts, e.g. *X°M$_{VT}$[3] is relevant in interrogatives, but not in declaratives, the HMC is vacuously satisfied in affirmatives, etc., by means of the superset/subset relation, as discussed in §3. Since grammatical principles invoked herein, such as economy of movement, are general across contexts, we can account for the fact that *do*-support is logically more prevalent in interrogatives than in declaratives. The same general principles apply in both cases, but the price of verb-raising is higher in interrogatives than in declaratives – thus, as verb-raising becomes more disfavored by the grammar in general (i.e. the promotion of the *X°M$_{VT}$ constraints), interrogatives display a consistently higher frequency of use of *do* through the EME period than do declaratives, which do not require as many violations of *X°M$_{VT}$.

In summary, a general reanalysis circa 1500, a reflex of the promotion of *X°M$_{VT}$, triggers a cataclysmic change, resulting in the creation of the modal class, the use of *to* as a marker of the infinitive and the beginning of the use of periphrastic *do*. The ranking of the *X°M$_{VT}$ constraint family fluctuates during the EME period, with an overall tendency towards higher rankings of *X°M$_{VT}$ as the period draws to a close, producing the general 'S'-curve pattern of *do* use. A later change, around the beginning of the 17th century, involving the promotion of FILL, disrupts this pattern somewhat, temporarily depressing the use of *do* in negative contexts and leading to the eventual obsolescence of periphrastic *do* in non-emphatic affirmative declaratives.

An Optimality-Theoretic account of language change allows one to describe the systematic replacement of one linguistic form by another through a systematic reranking of constraints (cf. Warner 1997), with periods of variation expressible as variability in constraint ranking (partial ordering of constraints). Furthermore, the interaction of constraints allows for the derivation of the relative proportion of occurrence of an innovation across contexts. Additionally, the interaction of constraints allows for an explication of deviations from systematic patterns of linguistic change – as observed in the case of *do*-support, because the occurrence of other changes – other constraint rerankings – may distort the surface pattern of an ongoing change. But the generality of the principles formally instantiated as *constraints* within the Optimality-Theoretic framework allows basic patterns of change, such as an increasing loss of verb-movement, to be understood in terms of basic changes in the priority of conflicting grammatical principles.

Johns Hopkins University

5. NOTES

[1] First and foremost I wish to thank Luigi Burzio—to whom I owe an enormous debt of gratitude for his skilled and patient supervision of this work since its conception—for his reading of, comments on and discussion of innumerable drafts, revisions and fragments. Without his contributions, suggestions, advice and general encouragement this work would never have reached fruition.

I should also like to thank Géraldine Legendre, Paul Smolensky and David Lightfoot for their insightful discussion of and expert advice on many aspects of this project. I am grateful as well to Paul Boersma, Bob Frank, Matt Goldrick, Paul Hagstrom, Chung-hye Han, Arja Nurmi, Bruce Tesar and Colin Wilson for suggestions, advice, assistance and/or discussion on various points.

I also appreciate the helpful comments from audiences who heard earlier versions of this work as presented at the Johns Hopkins University, the University of Maryland, College Park and at the Special Session on Optimality Theory and Language Change at the 45th Meeting of the International Linguistic Association held at Georgetown University—with special thanks to Eric Holt, for his organization of the ILA OT session, as well as his efforts in realizing this volume itself. My thanks to Luigi Burzio (once again) and to Jennifer Baggot for their assistance in bringing clarity to the presentation and exposition.

Thanks also to Paul Boersma, Larry LaFond, Eric Holt and two anonymous reviewers for corrections, helpful criticism and suggestions.

I alone bear responsibility for any remaining errors or spurious claims.

This work was supported by a Jacob K. Javits Fellowship [U.S. Dept. of Education] held by the author and an IGERT Grant [National Science Foundation] to the Johns Hopkins University Dept. of Cognitive Science.

[2] The proposed overlap in the principles, or constraints, of syntax and phonology does not imply that, for instance, the ranking of the FILL in phonology affects the ranking of the FILL constraint in syntax – the two modules are autonomous at least to that extent.

[3] Whether the SUBJECT constraint requires that the highest A-Spec position be filled (in which case the Italian example lacks any element in SpecIP) or whether it requires that the position be filled with on overt element (in which case pro occupies SpecIP), makes little difference for the focus of the chapter.

[4] Another possibility, explored in Zanuttini (1997) is that the lower position may be used in 'presuppositional negation' contexts. Both (7a) and (7b) are felicitous with a variety of counterfactual situations, e.g. the cat in question is still at home; the cat has gone instead to the Riviera on holiday; the cat has gone to London, but for a purpose other than seeing the Queen, etc. However, (7a) is more felicitous when it is uttered by a speaker in order to counter a discourse-presupposition that 'the cat has gone to London to see the Queen'.

[5] As this investigation does not concern the movement of XPs, we generally do not show traces relevant to such movement. The reader may assume that subjects are generated in SpecVP and are raised to satisfy an EPP or SUBJECT constraint (cf. Grimshaw 1997) – which demands that the highest A-specifier position of a clause be filled – and/or to satisfy case requirements. For the purposes of this paper it matters little what the reason for XP movement in these cases is considered to be (see Grimshaw 1997, Legendre et al. 1998 and Burzio 2000 for OT analyses of XP movement).

[6] The insertion of *do* into Mood° rather than Tense° avoids a violation of the undominated OBHD, as Mood° would otherwise be left empty – see below in main text.

[7] We assume, as is fairly standard, that *yes-no* questions involve a covert operator equivalent to a *wh*-word, relevantly in its need to take scope over the clause.

[8] This also provides a principled reason why *do*-support does not occur with subject *wh*-phrases – since a subject *wh*-phrase already has scope over the clause it need not move (assuming that A-positions are also valid operator positions), e.g. *Who wants to leave?* vs. **Who does want to leave?*. For further investigations into *wh*-movement in OT, see also Legendre et al. 1998, Ackema & Neeleman 1998.

[9] I assume that (bound) affixes that are not attached to a free-standing element (e.g. verbal head (lexical or expletive) or unbound inflection (= modal – see below)) cannot surface. Thus head-containing affixes left unattached are empty and therefore in violation of OBLIGATORY HEAD (cf. the 'Stranded Affix Filter' of Lasnik 1981).

[10] The pre-verbal *ne* is not used in colloquial French, and thus *pas* appears to be the 'primary' element of negation, in colloquial French at any rate, and is equivalent to English *not.*

[11] "Isolects are varieties of a language that differ only in a minimal way, say by the presence or weighting of a single feature in a rule, or by a minimal difference in rule ordering. A single isogloss stands between two isolects of a language" (Bailey 1973:11).

[12] This function is $p = e^{k+st} / 1 + e^{k+st}$, where p=fraction of the advancing form, t=time variable and s, k=constants; or the equivalent logistic transform of frequency: $p / (1-p) = k + st$, where the constant s = slope of the function, i.e. rate of replacement of the established form by the new form, k = the intercept parameter, which measures the frequency of the new form at the fixed point $t = 0$. There is no time t for which $p = 0$ or $p = 1$, but as t approaches $-\infty$, p approaches 0, and as t approaches ∞, p approaches one (Kroch 1989b:204).

[13] The possible ranking permutations of x constraints is $x!$. Thus, the total number of permutation of 4 constraints is $(4)(3)(2)(1)$=24. When constraints standing in universally-fixed rankings are included, in this case *X°MVT^3 » *X°MVT^2, then (simplifying a bit) one divides by the number of constraints that are universally ranked w.r.t. each other – thus $(4)(3)(2)(1)/(2)(1)$ gives us twelve rankings, as shown in (19).

[14] All examples in this section are drawn from Shakespeare's plays (Oxford 1988 edition with modernized spellings).

[15] Based on the analysis so far, one would expect auxiliaries to behave as regular verbs with respect to verb-raising, *do*-support, etc. However, this is not the case. Never in the history of English do auxiliaries display any other option than raising, i.e. we never find examples of the sort **John does not have left*. In order to guarantee the raising of auxiliaries notwithstanding the relative ranking of the *X°MVT constraints and FILL, PRBD, we posit the constraint that requires inflection to be attached to the auxiliary[15]:

V[+AUX]+INFL: An auxiliary verbal head must be attached to Agreement, Tense and Mood features.

This constraint is part of the V+INFL constraint family. A high ranking of this constraint will account for the non-occurrence of *do* with the auxiliaries throughout the history of English.

The fact that auxiliary-raising is blocked in the presence of modals, e.g. *John should be leaving* vs. **John is should leaving*, can be taken to reflect a lexicalization effect. That is, once the modals are established as a class, whenever the input contains information expressed by the modals (e.g. future *will*, epistemic-modality *could*, etc.) the grammar contains a 'parochial' constraint forcing such information to be expressed through the use of a modal, i.e. as an unbound inflection, such a constraint, higher-ranked than V[+AUX]+INFL, can be interpreted as imposing morphological invariance or 'anti-allomorphy' (Burzio 2000, to appear).

[16] Likewise for the other contexts, excepting the relation between affirmative interrogatives and negative declaratives, where a true superset-subset relation does not hold, because the of NegP1/NegP2 asymmetry – but the affirmative interrogatives still show a greater overall use of *do* than do negative declaratives.

[17] The status of Asp-to-C movement is irrelevant if the V-to-Asp operation is not active.

[18] Actually, there is category-raising of V to Asp (over the lower negation) with subsequent feature-movement to Mood and T, see Han & Kroch 2000 for details.

[19] Again, actually there is category-raising of V to Asp with subsequent feature-movement to Mood and T. As above, there is a question, on this account, why *do*-support *ever* appears in affirmative declarative contexts.

[20] In the remaining discussion we mainly set aside affix-lowering, focusing on the competition between verb-raising and *do*-support, in the interests of clarity, having discussed the complete competition between verb-raising, affix-lowering and *do*-support in §2 and §4 above.

6. TEXTUAL SOURCES

Bacon, Francis. *The Advancement of Learning* and *New Atlantis*. London: Oxford University Press, 1951.

Gray, Douglas, ed. 1985. *The Oxford Book of Late Medieval Prose and Verse.* Oxford: Clarendon Press.

Lamson, Roy and Hallett Smith, eds. 1942. *Renaissance England: Poetry and Prose from the Reformation to the Restoration.* New York: Norton.

Shakespeare, William. *The Complete Works.* S. Wells & G. Taylor, J. Jowett & W. Montgomery, eds. Oxford: Clarendon Press, 1988. [We gratefully acknowledge the use of Jeremy Hylton's Shakespeare web-site (http://the-tech.mit.edu/Shakespeare/) in carrying out textual searches. Quotations as per Oxford (1988) edition.]

As well many examples are taken from citations in the secondary sources. All examples thus cited are also referenced to the secondary source in which they appear.