Chapter 4

Do-Support and the Loss of Verb Movement

4.1. Introduction

English has a language-specific but highly regular operation, *do*-support. Though numerous attempts have so far been made on this operation through the history of generative grammar, there is little agreement as to how this phenomenon should be treated within the framework of the minimalist program.[1] One purpose of this chapter is to derive the language-specific nature of *do*-support from independent general principles by taking into consideration the historical change of verb movement. More specifically, it will be claimed that *do* is an X^0 expletive element whose primary function is to check off a strong feature on functional heads, so that *do*-support can be treated on a par with the insertion of the expletives *there* and *it*. It will also be shown that the distribution of the strong feature that triggers *do*-support in present-day English has resulted from the kind of diachronic process of feature shift we have seen in the preceding chapter.

From a theoretical point of view, an analysis of *do*-support will provide us with a solution to the problem concerning the driving force of diachronic feature shift. In the preceding chapter, it was demonstrated that when a feature diachronically shifts among functional heads, it can only appear at a neighboring head after a given parametric change (the locality property). At this point, however, it remains to be explained why such parametric changes occur in the first place. As argued in Chapter 1, as long as language acquisition is successfully carried out among successive generations, parameters would always be fixed to the same values and there would be no chance of language change. In this chapter, I will argue that the locality property of diachronic feature shift can be attributed to the morphological change of verbal affixes. To formulate properly effects of the decline of verbal morphology on parametric changes, I will revise the definition of *Attract F* again and introduce the notion of *Attraction Domain*, which says, in essence, that an attractor's procedure for identifying a matching feature must be executed within certain designated domains. This will open up a way of explaining parametric changes in terms of properties of UG.

The organization of this chapter is as follows. Section 4.2 reviews some basic issues and describes the distribution of the strong feature in present-day English. Section 4.3 examines the morphological structure of verbal inflections and its relevance to the checking theory. Then, section 4.4 advances proposals regarding the mechanism of verb movement. Section 4.5 is devoted to a historical

analysis of the loss of verb movement and the rise of *do*-support. On the basis of this historical analysis, section 4.6 presents an account of *do*-support in present-day English. Finally, section 4.7 deals with some residual problems.

4.2. Minimalist Issues of *Do*-Support

The advent of the checking theory in the minimalist model has brought about new problems concerning *do*-support and auxiliary raising. Let us review some major issues.

Chomsky (1995) assumes that the head Infl in English does not have a strong feature which induces overt verb movement. This explains the fact that English main verbs cannot move across VP adverbs, as opposed to French counterparts:

(1) a. *English*

John *quickly* learned his lines.

b. *French*

Jean apprendra *vite* son rôle.

(identical with (1a))

In (1a), Infl does not contain a strong feature, so that *learned* stays in situ in overt syntax; on the other hand, because French Infl contains a strong feature, the finite verb *apprendra* in (1b) appears to the left of the VP adverb *vite*. Even in English, however, the position of Infl or *c* must be filled with some auxiliary in the following environments:

(2) a. *negative sentences involving* not

Mary *did* not see John.

b. *emphatic sentences*

Mary *DID* see John.

c. *inverted interrogative sentences*

Who *did* Mary see?

When neither modals nor *have/be* are employed, *do* is obligatorily inserted as a dummy element. Even in affirmative declaratives, English auxiliaries must appear to the left of VP adverbs:

(3) a. Mary has *cheerfully* eaten the pizza.

b. Mary is *cheerfully* eating the pizza.

This indicates that these auxiliaries occur in a position higher than *v*P.

To capture these phenomena in terms of the checking theory, some proposals are made in the literature. Roberts (1998b), accepting the view that English Infl does not have a strong feature, proposes that the principle Procrastinate, which favors covert operations over overt ones, does not exist, and consequently all uninterpretable features are checked off before Spell-Out. This makes it possible that formal features of a finite verb alone move overtly, leaving its semantic and phonological features behind at the original position. Then he claims that auxiliary raising can be analyzed as overt feature-movement triggered by a weak feature of Infl. This analysis, however, crucially presupposes that auxiliaries are simply collections of formal features. This assumption is highly dubious since auxiliaries clearly have phonetic contents pronounced such as *have* and *be*, in addition to their formal properties.[2] Thus,

it can be said that the trigger of *have/be* raising and *do*-support must be a strong feature which requires pied-piping of the whole category.

Arimura (1998) gives a somewhat different characterization to overt feature-movement. He proposes that English Infl uniformly has a strong V feature and overt movement it induces carries along the whole category only when the derivation converges at LF. He further assumes that θ-roles are assigned at LF and English Infl is opaque with respect to θ-role assignment. If a finite main verb moves to Infl with its whole category in overt syntax, it will cause a violation of the θ-criterion, because θ-roles cannot be properly assigned from that position; thus, only its formal features can be attracted by the strong V feature. On the other hand, nothing blocks pied-piping of the whole category in the case of auxiliaries, which lack θ-roles to assign. Therefore, an auxiliary moves to Infl along with its phonological features and is pronounced at the landing site. This analysis, however, still faces some problems. First, under Arimura's assumptions, at the stage where Infl attracts a finite verb, application of pied-piping must be determined on the basis of information about the final stage of the derivation; in other words, the attractor must know in advance whether or not pied-piping will lead to convergence at LF. This kind of "look-ahead" property should be avoided on conceptual grounds for reasons of reduction of computational complexity. A second problem is the status of the θ-theory. If semantic and phonological features of main verbs must stay in situ for reasons of θ-role assignment, the same would hold for

argument DPs as well. Since Arimura's strong features do not necessarily require pied-piping of the whole category, the EPP property of Infl would be satisfied by formal features of the subject alone, with its semantic and phonological features remaining in situ to receive its θ-role, but this is of course contrary to fact.

It seems to me that all these problems arise from positing the same strength of features irrespective of sentence types, whether it is strong or weak. Roberts sticks to the hypothesis that the V feature of Infl is always weak in English and fails to account for auxiliary raising. Arimura is forced to make stipulative assumptions concerning θ-role assignment because he starts off from the position that the V feature of Infl is always strong. However, there is no a priori reason to rule out the possibility that the distribution of the strong feature in affirmative declaratives is different from that in sentences where Infl or *c* is obligatorily filled with some auxiliary. Such a heterogeneous distribution of a strong feature may not be preferable in view of simplicity of the grammatical theory, but we will nevertheless pursue this line of inquiry and show that this can be ultimately reduced to properties of UG if we extend our observation into the history of verb movement.

An argument for the hypothesis that *do*-support is triggered by a strong feature comes from striking similarity between the dummy *do* and the expletives *there* and *it*, as illustrated in (4):

(4) a. [$_{IP}$ there$_i$ Infl{~~strong D~~} seems [$_{IP}$ t_i Infl{~~strong D~~} to be someone in the room]]

b. [$_{IP}$ it Infl{~~strong D~~} seems [that someone is in the room]]

The expletives *there* and *it* share a number of syntactic properties with the dummy *do*: these elements all lack semantic contents such as θ-roles and thus are inserted directly into the domain of functional categories without being base-generated in VP-internal positions. Following the standard view of Chomsky (1995), let us assume that *there* and *it* in (4) are inserted into [Spec, IP] to check off the strong D feature of Infl and satisfy the EPP. If this is correct, then it will be reasonable to suppose that in the environments where the dummy *do* is inserted, Infl or *c* has a certain strong feature which must be checked off by the whole category of V before Spell-Out. This is to say that *do* is the X^0-counterpart of the expletives *there* and *it*, which are primarily employed to satisfy a strong feature of a functional head.

Then, to ensure that the dummy *do* will be inserted in negative sentences, emphatic sentences and matrix interrogatives and that *have/be* raising takes place overtly, we must posit the following distribution of a strong feature. Following the discussion in Chapter 3, let us assume that the relevant feature is [+F], which attracts an element with a [+finite] feature:[3]

(5) a. In affirmative declaratives, matrix *wh*-subject interrogatives, and embedded interrogatives in general, [+F] is located no higher than *v*.

b. In negative sentences, emphatic sentences, and sentences including auxiliaries, [+F] is located in Infl.

c. In matrix inverted interrogatives, [+F] is located in *c*.

As for matrix inverted interrogatives, Chomsky (1995:290) claims

that the strong Q feature of C (= *c*, in our terms) triggers auxiliary raising, but this analysis leaves us with an unsolved problem. For Chomsky, the Q feature of *c* is checked off either by X^0-adjunction to *c*, as in (6a), or by XP-insertion to [Spec, *c*P], as in (6b):

(6) a. did+Q [$_{IP}$ John give a book to Mary]

b. I wonder [$_{cP}$ which book$_i$ Q [$_{IP}$ John gave t_i to Mary]]

c. which book$_i$ did+Q [$_{IP}$ John give t_i to Mary]

However, a redundancy arises in sentences like (6c) in which both a *wh*-phrase and an auxiliary undergo movement. If Q is checked off against *which book*, auxiliary raising will not be necessary; on the other hand, if Q is checked off by adjunction of *did*, nothing will motivate the movement of the *wh*-phrase.[4]

Bobaljik (1995) puts forward a hypothesis that *do*-support is a PF operation that is applied when Infl is not adjacent to the verbal stem in the phonological component. It is argued that in negative sentences, *do* is inserted to support Infl because the negative marker *not* blocks adjacency between V and Infl; likewise in inverted interrogatives, a lexical subject in [Spec, IP] blocks the same kind of adjacency given that *wh*-movement to [Spec, Comp] has triggered inversion of Infl to Comp. However, it remains unexplained in his analysis why Infl raises to Comp in the first place in inverted interrogatives. Without Infl-to-Comp movement in overt syntax, *do*-support would be unnecessary because there would be no elements intervening between Infl and V. Therefore, the PF-insertion approach also fails to account for the distribution of *do*

correctly. Then, if we analyze *do*-support as a purely syntactic operation without recourse to properties of phonological component, it will be inevitable to posit an independent feature like [+F] and assume the distribution in (5).

At this point, (5) merely paraphrases the linguistic facts about *do*-support and explains nothing. Then the problem at issue is why [+F] exhibits such an irregular distribution in different types of sentences in present-day English. In section 4.5, I will demonstrate that (5) is a natural consequence of the loss of verb movement in the history of English, but before going to that, let us examine another phenomenon related to the auxiliary/main verb asymmetry.

4.3. Verbal Morphology and Overt V-Raising

In this section, I will first argue that the auxiliary/main verb asymmetry in English is a reflection of their difference in morphological structures and then show that the richness of main verb inflections has much to do with the possibility of overt V-raising.

4.3.1. The Locus of the [+finite] Feature

Let us start with Warner's (1993) observation about VP ellipsis in English. He argues that there is sharp contrast between ellipses headed by auxiliaries and main verbs. When an elided VP is headed by an auxiliary verb, the auxiliary must have the very same morphological form as its overt antecedent, as in (7) (elided VPs and their antecedents are indicated by solid lines and square brackets,

respectively):

(7) a. John will always [be a brute], and I expect his son will ~~be a brute~~ too.

b. If John was [being awkward] yesterday, no doubt he is ~~being awkward~~ again today. (Warner (1993:53))

If two VPs contain morphologically distinct auxiliaries, deleting one of them greatly reduces acceptability of the sentence. This is illustrated in (8):

(8) a. *Paula was [being really wilful] this evening. I do wish she wouldn't ~~be really wilful~~.

b. *John has already [been unkind to Ruth] for most of the afternoon, so at this very moment he probably still is ~~being unkind to Ruth~~.

c. *The lower floors [haven't caught fire], so the upper floors can't ~~have caught fire~~ yet either.

d.?*You [haven't seen one] yet? You should ~~have seen one~~ by now if they're really there. (ibid.:54-55)

Morphological discrepancy between the two *be*-phrases is obvious in (8a, b). In addition, the examples in (8c, d) indicate that the non-finite *have* cannot be deleted when its (morphologically identical) antecedent is used as a finite form.[5]

On the other hand, in the case of ellipsis of VP headed by a main verb, strict morphological identity does not seem to be required. Compare the examples below with those in (8):

(9) a. Paula was [making snide remarks] all afternoon. I do wish she wouldn't ~~make snide remarks~~.

b. ?John has already [teased Ruth] for most of the afternoon, so at this very moment he probably still is ~~teasing Ruth~~.

c. Max is [selling hot dogs for a living], and soon all of us will have to ~~sell hot dogs for a living~~.

d. John said he would never [take money on the side], but I knew he was ~~taking money on the side~~.

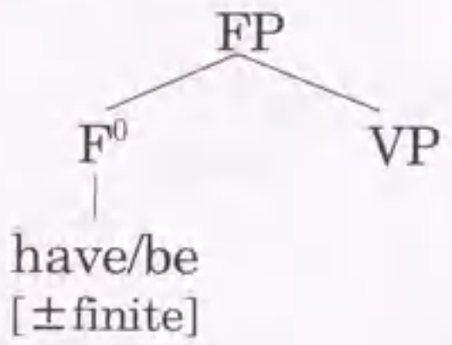
(ibid.:50, 51, 54)

For example, (9a) is acceptable even though the two non-finite forms of *make*, present participle and infinitive, are not identical. This shows a sharp contrast with the ellipses of the auxiliary phrases in (8).

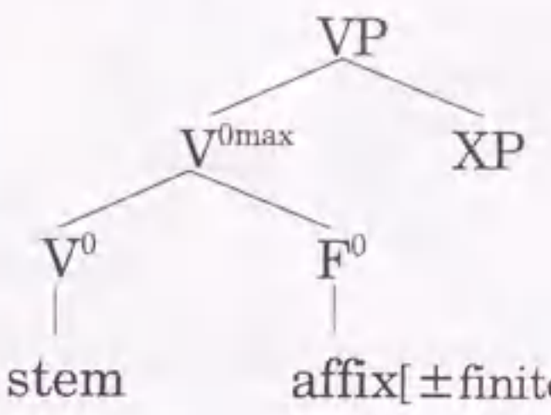
It appears that the sameness condition on ellipsis applies only to auxiliary phrases, not to main verb phrases. To reconcile the data above with the simplest and conceptually the most desirable hypothesis that VP can only be deleted under strict identity with the antecedent, Lasnik (1995) proposes that auxiliaries are always thrown into phrase structures fully inflected whereas stems and affixes of main verbs separately enter the computational system. If stems of main verbs are structurally independent of affixes and VP ellipsis is carried out before these two elements merge, the apparent violation of the sameness condition will turn out to be harmless to the strong hypothesis.

Essentially following Lasnik's proposal, let us suppose that auxiliaries and main verbs have different internal structures. In the present general framework, we can delineate them as follows:

(10) a. *Auxiliaries*



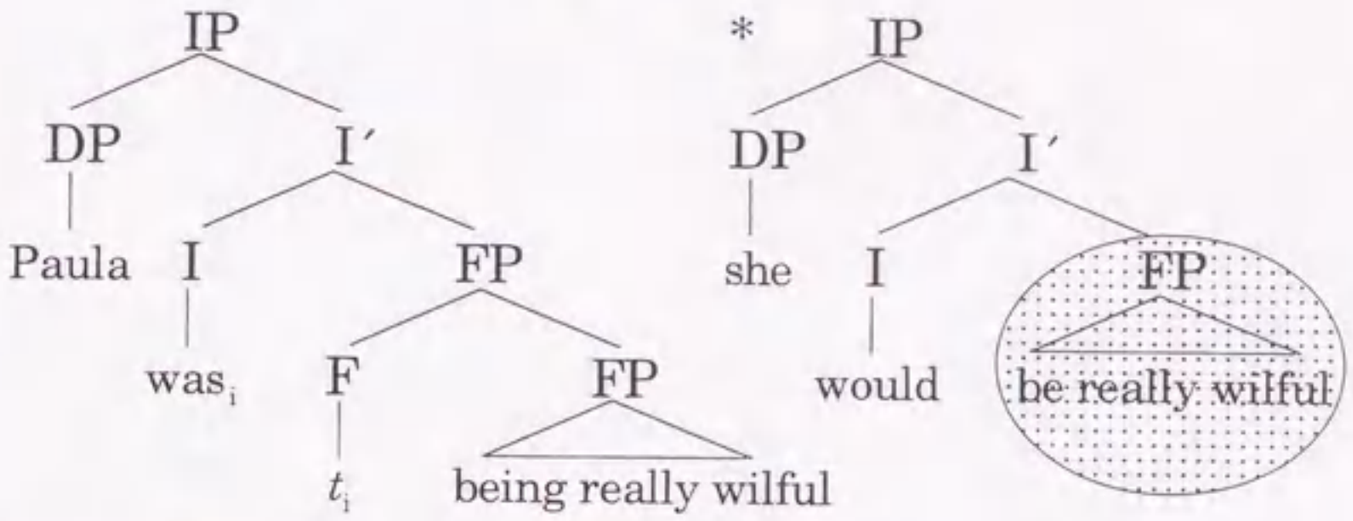
b. *Main Verbs*



Let F^0 be a functional category that bears the [+finite] feature for finite verbs and the [−finite] feature for non-finite verbs. While auxiliaries *have/be* are themselves dominated by F^0, in the case of main verbs F^0 dominates only affixes. Though Lasnik originally proposes that a verbal affix takes VP headed by a stem as its complement and the two elements undergo PF merger, we assume that F^0 in (10b) directly adjoins to the stem by Merge in the overt syntax. This yields the structure in which the [+finite] feature is deeply embedded under V^{0max}. In other words, the affix in (10b) is a head that does not project any further. The contrast between (10a) and (10b) concerning the "depth" of the [+finite] feature will play a crucial role in our later explanation for the auxiliary/main verb asymmetry reviewed in the preceding section.

It immediately follows from the structural difference that only ellipses of auxiliary phrases apparently require strict identity with antecedents. The relevant part of (8a) can be represented as in (11) (the shaded portion stands for the target of ellipsis):

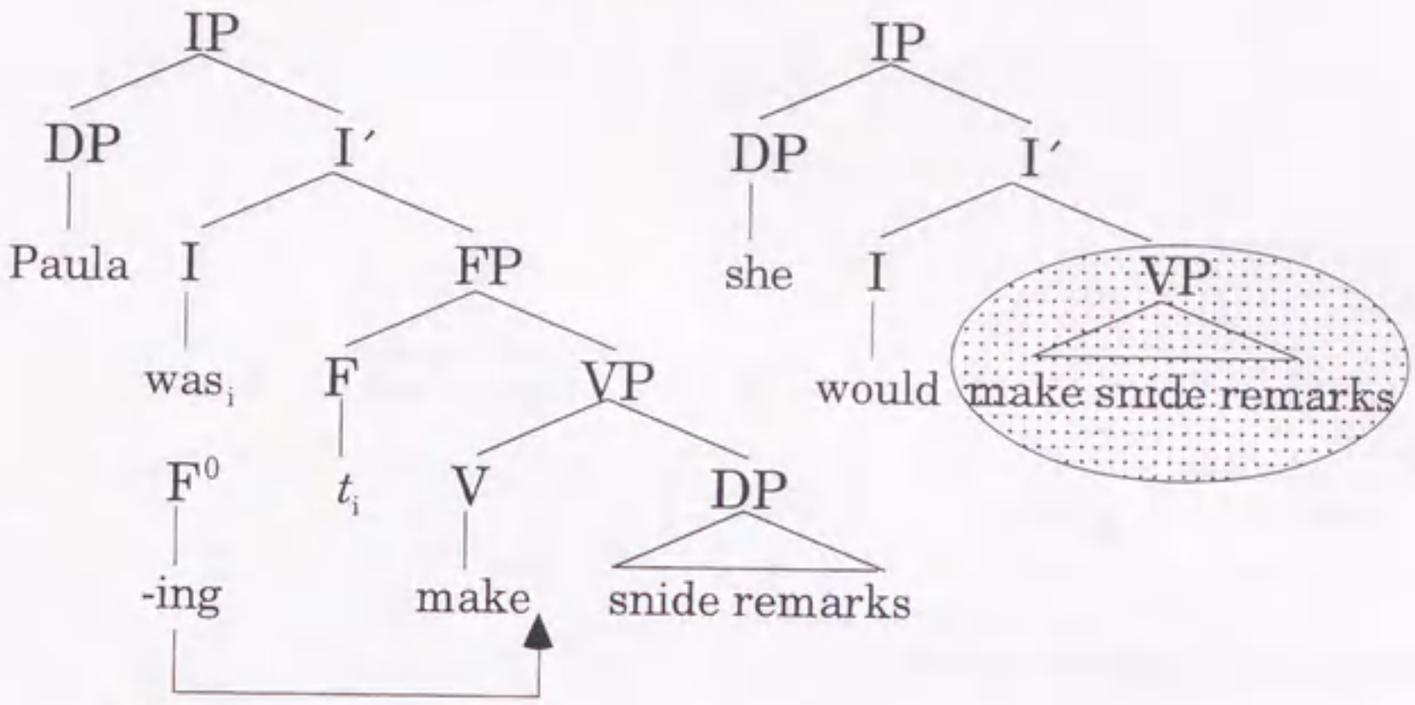
(11) *Paula was [being really wilful] this evening. I do wish she wouldn't ~~be really wilful~~.



Suppose that ellipsis is an instance of a syntactic operation *Delete*, which renders some features invisible at one of the interface levels. In the cases under discussion, Delete makes some phonological features invisible at PF and possibly syntactic features invisible at LF under strict identity with the antecedent (i.e. the sameness condition). Since auxiliaries enter the computational system fully inflected, the target FP in (11), whose head has different form from its antecedent, cannot be deleted.

Main verbs are also subject to the sameness condition on ellipsis, but the result is not reflected at the PF output. (12) is the relevant structure of (9a):

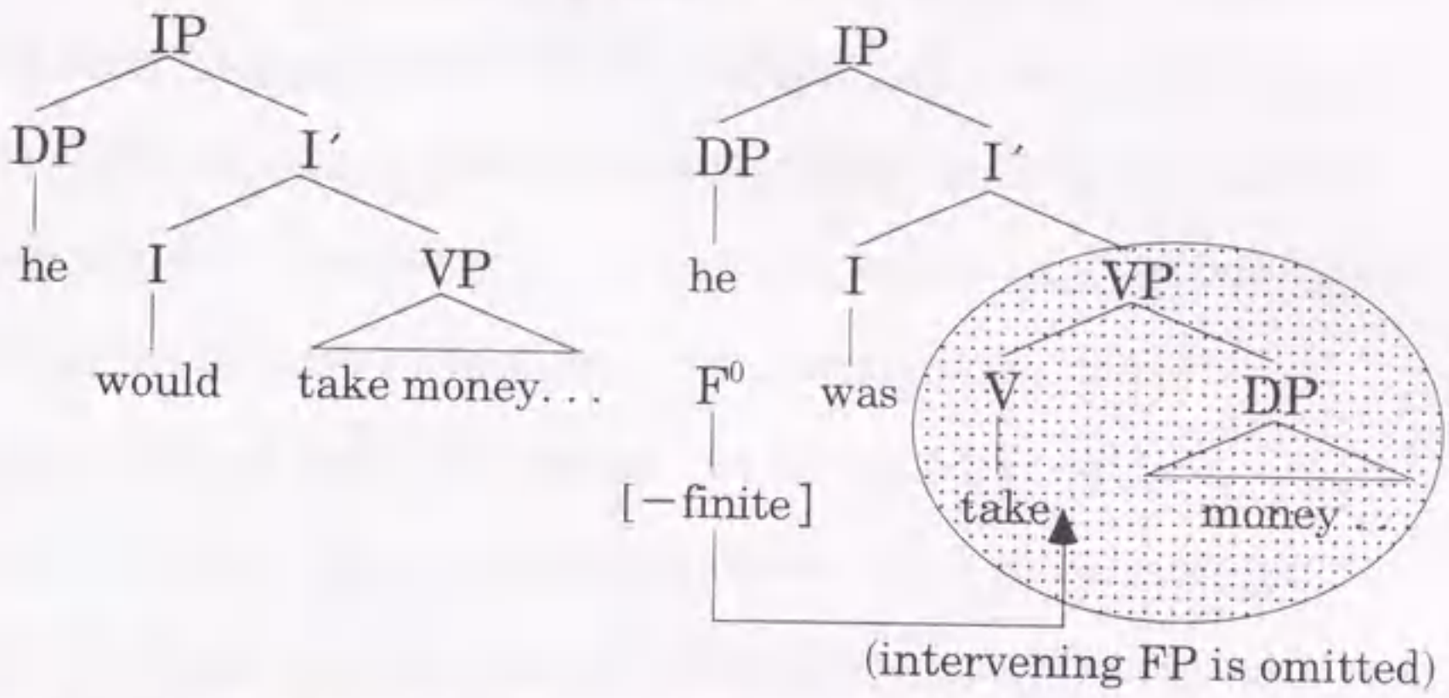
(12) Paula was [making snide remarks] all afternoon. I do wish she wouldn't ~~make snide remarks~~.



Affixes of main verbs are thrown into phrase structures independently of their stems, so that it is possible that Delete applies to the target VP in accordance with the sameness condition before the affix dominated by F^0 is attached to the head of the antecedent VP. As a result, the antecedent assumes an inflectional affix distinct from the target.[6]

Cases like (9d), repeated here as (13), can also be accounted for by exactly the same reasoning:

(13) John said he would never [take money on the side], but I knew he was ~~taking money on the side~~.



In this example, the head of the target VP, not the antecedent, has an inflectional affix. As in (12), phonological features of the target VP are deleted before Spell-Out under identity with the antecedent. After that, only formal and semantic features of F^0 including the [−finite] feature adjoin to the stem in the covert syntax. As Chomsky (1995:292) suggests, nothing in principle prevents covert merger of a lexical item that has no phonological features. Consequently, the introduction of F^0 in (13) has no PF effect.[7]

Given the discussion of VP (or FP) ellipsis in this section, it can be concluded that we have obtained independent evidence for the verbal structures in (10).

4.3.2. Richness of Inflection

To see a little more fully the relevance of verbal morphology to overt verb movement, let us pay attention to the richness of inflectional affixes. It has been often pointed out that the presence or

absence of obligatory verb raising to Infl is closely connected with the richness of verbal inflectional morphology. For example, Pollock (1989) claims that in languages like French which have relatively rich verbal inflections, Agr (=Infl) is transparent and thus allows a main verb adjoined to it to head a θ-chain, whereas in English Agr is opaque because of its poor inflection, and raising a main verb to it will yield a violation of the θ-criterion. In the more recent framework of Chomsky (1995), however, overt verb movement is abstracted away from the richness of inflectional morphology and it is simply reduced to the presence of a strong V feature on Infl, which calls for the category V in its checking domain, not tense or ϕ-features (Chomsky (1995:232)). Nevertheless, the correlation between overt verb movement and rich verb inflections is so widely observed that it is likely that this reflects some significant property of UG. Vikner (1997) investigates French and seven Germanic languages (English, Danish, Hallingmålet (Norwegian dialect), Faroese, Älvdalsmålet (Swedish dialect), Icelandic, Yiddish) and draws the following generalization:

(14) An SVO language has V-to-I movement if and only if person morphology is found in all tenses.

(Vikner (1997:200))

Vikner limits his discussion to SVO languages, but German and Dutch, whose base word order is SOV, also satisfy this criterion (see (17) below).[8] In any case, to the extent that this generalization holds true, the correlation must be accounted for in some way or

other regardless of theoretical frameworks.

Interestingly enough, (14) also holds for the loss of V-to-I movement in the history of English. Consider the paradigms in (15) for London-area English around 1500 and 1575 (cf. Gray (1985:495f), Roberts (1993:257), Vikner (1997:203), Araki and Ukaji (1984:196-202)):

(15)

1500 (with V-to-I)

	Present	Past
Sg. 1	- ϕ	-ed
2	-est	-ed-**est**
3	-eth	-ed
Pl. 1	- ϕ	-ed
2	- ϕ	-ed
3	- ϕ	-ed

1575 (without V-to-I)

	Present	Past
Sg. 1	- ϕ	-ed
2	-est	-ed-**(est)**
3	-eth	-ed
Pl. 1	- ϕ	-ed
2	- ϕ	-ed
3	- ϕ	-ed

V-to-I movement was still productive around 1500, and as expected, person morphology is found in both present and past tenses: *-est* and *-eth* for present second and third singular and *-est* for past second singular. Later around 1575, V-to-I was lost in English (Roberts (1993) and Watanabe (1994) based on data provided by Ellegård (1953) and Kroch (1989)). Vikner accepts the view of Pyles (1964:205) and Strang (1970:203) that the personal ending *-est* for past second singular began to be lost in the 16th century. If this is correct, it can be said that there was no distinct person morphology in past tense by 1575. Therefore, the loss of V-to-I movement in

English is straightforwardly captured by the generalization in (14).[9]

Here, I would like to take a step forward and draw attention to the relevance of rich verbal inflections to another kind of head movement: namely, V-to-*c* movement (V2). From OE through early ME, when V2 was prevalent, verbal morphology was even richer than in 1500, as illustrated below (cf. Ono and Nakao (1980: 261-273), Nakao (1972: 156-162)):

(16)

OE (with V2)

	Present	Past
Sg. 1	-e	-ed-e
2	-est	-ed-est
3	-eþ	-ed-e
Pl. 1	-aþ	-ed-on
2	-aþ	-ed-on
3	-aþ	-ed-on

(weak verb class I)

Early ME (with V2)

	Present	Past
Sg. 1	-e	-d-e
2	-est	-d-est
3	-eþ	-d-e
Pl. 1	-eþ/en	-d-en
2	-eþ/en	-d-en
3	-eþ/en	-d-en

(weak verb class I)

In OE and early ME, agreement morphology was completely preserved in both present and past tenses, and past tense was distinctively marked from present tense by the tense morpheme -(*e*)*d*. I will argue that the erosion of this very rich verbal morphology caused the loss of V2 in English in late ME in the same way as the loss of V-to-I movement in the 16th century.

Other Germanic V2 languages (German, Dutch, and Icelandic) have verbal inflections which are as rich as OE and early ME. Compare their paradigms for regular verbs with those of Romance

languages (French, Spanish, and Italian), in which finite verbs raise only to the Infl position:

(17) *Languages with V-to-*c *Movement*

	German		*Dutch*		*Icelandic*	
	Present	Past	Present	Past	Present	Past
Sg. 1	-e	**-t-e**	- ϕ	**-t-e**	- ϕ	- **ð-i**
2	-st	**-t-est**	-t	**-t-e**	-r	- **ð-ir**
3	-t	**-t-e**	-t	**-t-e**	-r	- **ð-i**
Pl. 1	-en	**-t-en**	-en	**-t-en**	-um	- **ð-u m**
2	-t	**-t-et**	-t	**-t-e(n)**	-ið	- **ð-uð**
3	-en	**-t-en**	-en	**-t-en**	- ϕ	- **ð-u**

(18) *Languages with V-to-I movement*

	French		*Spanish*		*Italian*	
	Present	Past	Present	Past	Present	Past
Sg. 1	-e	**-ai**	-o	**-é**	-o	**-ai**
2	-es	**-as**	-as	**-aste**	-i	**-asti**
3	-e	**-a**	-a	**-ó**	-a	**-o**
Pl. 1	-ons	**-âmes**	-amos	**-amos**	-iamo	**-ammo**
2	-ez	**-âtes**	-áis	**-asteis**	-ate	**-aste**
3	-ent	**-erent**	-an	**-aron**	-ano	**-arono**

A crucial difference is found in past tense forms. In Germanic V2 languages, there are distinct tense morphemes (*-t* and *-ð*) to which agreement morphemes are attached. In Romance languages, on the

other hand, there is only one slot for suffixes, which represent both tense and agreement at the same time. Though it is true that Romance languages have many different inflectional forms, we can reasonably say that Germanic verbal morphology is richer than that of Romance languages in terms of the number of morphemes that follow the stem.[10]

Then, I would draw the following generalization about V-to-*c* movement:

(19) A language has V-to-*c* movement if both tense and agreement morphemes are found in all past tense forms.

Germanic V2 languages including OE and early ME fulfill this criterion. Verbal inflections of Romance languages and early Modern English are not rich enough to satisfy (19), so that they have only V-to-I movement. It is worth pointing out here that in contrast to Vikner's generalization about V-to-I movement, (19) is not a necessary condition on V-to-*c* movement. This is because there are in fact languages with flat morphology that exhibit V-to-*c* movement. A representative case that immediately comes to mind is mainland Scandinavian languages (Swedish, Danish, and Norwegian). These languages prevent us from directly associating V2 with rich morphology, but I will nevertheless demonstrate below that if they are put aside for the moment, (19) works quite well in explaining the loss of V2 in the history of English. As for the treatment of mainland Scandinavian languages, we will return in section 4.7.

Historical facts confirm this generalization. It is assumed in

the literature (cf. van Kemenade (1987, 1997)) that V2 virtually died out around 1400 or in the first half of the 15th century. At the same time, Nakao (1972:161) observes that in midland and southern areas, the agreement morpheme *-en* for past plural forms was lost by the end of the 14th century. To see that these two simultaneous changes are not mere coincidence, I examined *Mandeville's Travels* and *The Paston Letters* from the first half of the 15th century. These texts were both written in the transitional period where V2 rapidly decreased, but they show us very different aspects of verbal inflections and word order. In *Mandeville's Travels*, inflectional morphology is well preserved and both tense and agreement morphemes are found in past plural forms, as shown in (20):

(20) a. þei *founden* .iij. crosses
they found three crosses
(*Mandeville's Travels* 51.20)

b. þei *fayleden* of hire purpos
they failed of their purpose (ibid. 126.19)

As expected from (19), this text has abundant examples of the Topic-V-NP subject word order, which can be derived only in V2 languages. Some of the examples are given below:

(21) a. And from this cytee brought Sampson the stronge the
and from this city brought Sampson the strong the
ȝates vpon an high lond
gates upon a highland (ibid. 20.9-10)

b. And þorgh þat flom passeden the children of Israel
and through that river crossed the children of Israel

all drye feet
all dry feet (ibid. 69.23-24)

In *The Paston Letters*, on the other hand, number and person are not distinctively marked in past plural forms and only the tense morpheme *-d(e)* appears:

(22) a. thei *seide* with a good herte
they said with a good heart
(1444 *The Paston Letters* 56.49-50)

b. thei *herd* tydynges that the freyr was comyng
they heard tidings that the friend was coming
(1443 ibid. 50.3-4)

Interestingly enough, when the sentence initial position is occupied by a topic element and the subject is a full NP, the subject, but not the finite verb, follows the topic:

(23) a. This day at x. of the clok Edmund Paston and the
this day at ten of the clock Edmund Paston and the
parson of Oxened went owth of the Manor doun to
person of Oxened went out of the manor down to
Wantown Gapp,
Wantown Gapp (1443 ibid. 50.1-3)

b. And with in a sevenygt after my master Berney sent
and within a seven night after my master Berney sent
Davy to my mastres
Davy to my mistress (1448 ibid. 81.42-43)

These examples clearly indicate that the finite V raises no farther than the Infl position. Thus, the fact that the unambiguous V2 word order is observed only in the text with relatively rich verbal

morphology provides a substantial support for our hypothesis that the demise of the past plural morpheme *-en* caused the loss of V2.

To sum up, I suggest that the familiar dichotomy of rich/poor (or strong/weak) inflection is not sufficient for a unified theory of verb movement. Instead, it is necessary to classify the richness of verbal morphology into the following three degrees:

(24) Degree I: Flat morphology, or person morphology is absent in some tense.

⇨ *Only V-to-*v *movement*

Degree II: Person morphology is found in all tenses.

⇨ *V-to-I movement*

Degree III: Both tense and agreement morphemes are found in all past tense forms.

⇨ *V-to-*c *movement*

To put it informally, the distance of verb movement is proportional to the richness of verbal morphology: the richer the inflection, the longer verbs can move. Having made this classification, we can now go on to consider how to incorporate this general tendency into the mechanism of Attract F and the checking theory. This is our task in the next section.

4.4. Proposals

We have seen above that (i) the [+finite] feature of main verbs is located in inflectional affixes, which is deeply embedded under V^{0max} and (ii) the distance of overt verb movement is determined by

the degree of richness of verbal morphology. To capture these points within the theoretical framework adopted here, I will introduce two novel notions, *Attraction Domain* and *Feature Collaboration*.

4.4.1. Attraction Domain

First, I would like to revise the definition of Attract F again in the following way:

(25) *Attract F* (final version)

a. An attractor K identifies the closest feature that can enter into a checking relation with a sublabel of K, iff K is in the *Attraction Domain* of the feature.

b. A feature identified by K moves to K through every possible landing site:

(i) A head position for head movement

(ii) Spec of IP for A-movement

(iii) Spec of Comp and *v*P for A-bar movement

The Attraction Domain is defined as in (26):[11]

(26) a. The Attraction Domain of a feature on X^{0max} or XP is the root.

b. The Attraction Domain of a feature on an affixal F^0 is

(i) *v*P if F^0 is a Degree I verbal affix.

(ii) IP if F^0 is a Degree II verbal affix.

(iii) *c*P if F^0 is a Degree III verbal affix.

The earlier version of Attract F (cf. (35) in Chapter 2) postulated that however distant an attractor K is from a matching feature *f*, K can

identify *f* unless there is a feature of the same kind between K and *f*. As for identification of a feature on X^{0max} or XP, little need to be said about the validity of this assumption. Successive cyclic *wh*-movement as in (27) affords a good illustration of this point:

(27) $[_{cP}$ what$_i$ do+*c* [+wh] you $[_{vP}\ t_i$ think $[_{cP}\ t_i$ that [−wh] John $[_{vP}\ t_i$ said $[_{cP}\ t_i$ that [−wh] you have eaten $t_i]]]]]$

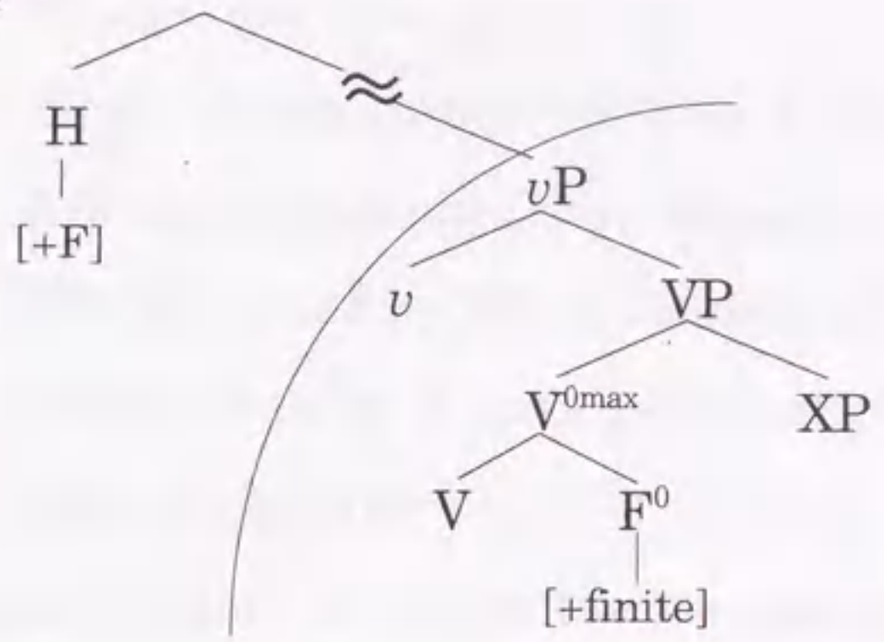
The [+wh] feature of the matrix *c* identifies the *wh*-phrase *what*, which is generated as the complement of the verb *eaten* in the lowest *c*P, and then *what* moves to the target position in accordance with the relativized minimality because of clause (25biii). Thus, (26a) amounts to saying that there are virtually no restrictions imposed on the distance between K and *f* insofar as *f* is on X^{0max} or XP.

The contrast between (26a) and (26b) is intended to capture the difference in internal morphological structures between auxiliaries and main verbs. Since auxiliaries have no internal branching and bear the [+finite] feature by themselves, this feature is always visible to identification by [+F]. By contrast, a [+finite] feature which resides in the morphological level (i.e. F^0 dominated by V^{0max}) must be sufficiently rich in order to be visible to long-distance identification and to move to the target position carrying along the stem. (26b) prescribes the correlations between the richness of verbal inflections and the domains in which affixes are visible to the attractor.

For example, suppose that the [+finite] feature is carried by a Degree I affix and [+F] is located in a functional category H higher

than *v*P:

(28) *



The arc stands for the Attraction Domain of the Degree I affix. The attractor [+F] and the attractee [+finite] can be compared to a telescope and a light source, respectively. To put it metaphorically, a poor affix can emit only weak light, so that [+F], a telescope, cannot find it in the far distance, as in (28). By contrast, if an affix is rich enough to emit strong light that covers the area in which [+F] is located, [+F] successfully identifies the [+finite] feature and attracts it to its checking domain. Thus, the insight behind the definition of the Attraction Domain is that a rich inflection can make itself seen in a wider area than a poor inflection.

4.4.2. *Feature Collaboration*

Feature identification depends upon not only the Attraction Domain of an attractee but also the property of an attractor. I propose that an attractor sometimes conspires with another feature to identify its attractee. I define this kind of *Feature Collaboration*

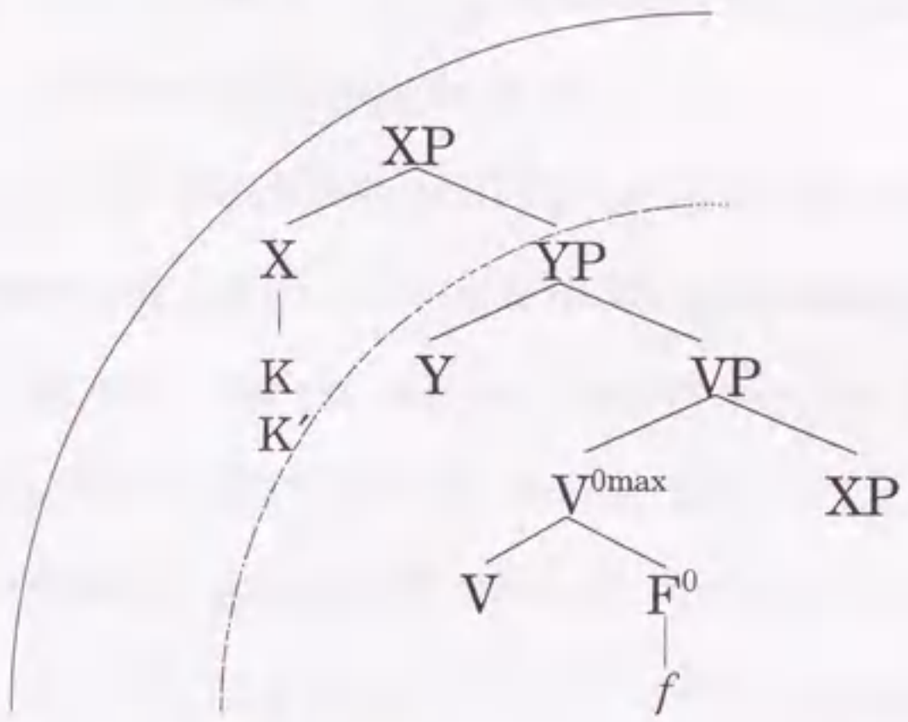
as follows:

(29) *Feature Collaboration*

When an attractor K is on the head of XP and the Attraction Domain of the closest matching feature *f* is YP dominated by XP, K identifies *f* iff K is supported by another potential attractor K′ and there is no ZP between XP and YP.

The tree representation for (29) is the following:

(30)



Even when K is outside the Attraction Domain of an inflectional affix *f* which has the strength of Degree α, once K is supported by another feature K′, *f* behaves as if it had the strength of Degree $\alpha+1$, so that identification of *f* by K is properly implemented insofar as there are no maximal projections intervening between XP and YP in (30). To use the metaphor again, even if a light source emits only weak light, a telescope can identify it when it is powerful enough to attest such weak light. Though K′ itself is not checked off, it helps the attractor K to identify a distant matching feature *f*, serving as a

kind of catalyst.

Of course, it is not the case that K can be supported by any K′. For example, if a weak V feature of Infl in present-day English could collaborate with ϕ-features, the weak feature would be strengthened so that overt V-to-I movement would occur freely. Then, to more narrowly define possible combinations of features that can enter into feature collaboration, I add the condition in (31) to the definition above:

(31) An attractor K or another potential attractor K′ must be an operator feature.

I expect that this condition will be ultimately derived from some other independent principle in a more articulated feature theory but leave this question for future research. For the present concern, I propose that the ability of [+F] to identify a [+finite] feature can be strengthened with the aid of a [+wh] feature of *c* on the following grounds. Let us first assume with Holmberg and Platzack (1995) that *finiteness* is a concept that is prerequisite for tense and mood and guarantees the existence of predication at the time of utterances. We then adopt Pollock's (1989) quantification theory of verb movement and assume that [+F] serves as an operator.[12] By attracting an element with a [+finite] feature (i.e. a finite verb), [+F] forms an operator-variable configuration (V . . . *t*) and thereby licences the finiteness of sentences; in other words, a state of affairs expressed by a finite verb is mapped into some point of the time axis by means of [+F]. Given this characterization of [+F], the

collaboration of [+F] with a [+wh] feature is not surprising, since a [+wh] feature also counts as an operator that takes scope at LF and therefore both of these features satisfy the condition in (31).

The idea that an attractor supported by another feature can induce otherwise illegitimate movement is independently put forth by some studies. Consider, for example, the following inversion constructions:

(32) a. *Locative Inversion*

Into the room entered the man.

b. *Quotative Inversion*

"I am so happy," thought Mary.

In these constructions, the finite verbs appear to the left of the subjects, indicating that V must have raised to a position higher than *v*P. In his study of locative inversion, Wakayama (1996) assumes a structure in which locative PP occupies [Spec, IP], V raises to Infl, and the subject remains in the base-generated position. He then argues that this V movement is induced by the [Agr] feature of Infl supported by a [+Loc] feature. According to his analysis, English [Agr] is not strong enough to attract V by itself, but thanks to the help of [+Loc], V can raise overtly to Infl. Collins (1997) also proposes a similar structure for quotative inversion in (32b). The only difference is that not the quote itself but a null quotative operator occupies [Spec, IP] and the quote is coindexed with it. He assumes that this quotative operator has a D feature which checks off the EPP feature of Infl. Then, to account for overt V movement in

quotative inversion, he proposes that "the EPP feature of T may enter into a checking relation with the quotative operator only if V[Quote] adjoins to T" (p. 41). This means that the EPP feature of T can attract the quotative operator only if it is supported by the categorial V feature. In our terms, Wakayama's [+Loc] feature can be regarded as a kind of operator feature; thus, the condition on feature collaboration in (31) is satisfied in this case. Collins's analysis of quotative inversion also involves operator movement, so that it is very likely that his proposal concerning exceptional verb movement can be restated in terms of a more general system of feature collaboration, though we will not go into details here.

4.5. Explaining the Distribution of [+F]

Having laid out the theoretical framework, let us return to the question raised in section 4.2. Present-day English exhibits a seemingly arbitrary distribution of [+F]: it is located in *v* in affirmative declaratives, in Infl in negative and emphatic sentences, and in *c* in inverted interrogatives. I will demonstrate that this irregular distribution has resulted from the diachronic feature shift of [+F] caused by the demise of inflectional affixes of finite verbs. The definition of Attract F which incorporates the notion of the Attraction Domain, together with the mechanism of feature collaboration, will enable us to account for parameter resetting solely by properties of UG and a widely accepted assumption about language acquisition in (33):

(33) The language learner chooses parametric values that maximally reproduce the surface pattern of primary linguistic data (PLD).

In what follows, we will first explain the loss of V2 in the mid-15th century and then go on to the loss of V-to-I and the rise of *do*-support in the 16th century.

4.5.1. The Loss of V2

In order to examine the change in the mid-15th century, we must first consider how children set the V2 parameter in OE. They encounter the following surface word order patterns when they acquire the grammar of OE (the examples are taken from van Kemenade (1987)):

(34) a. *topic-V-NP subject*

maran cyðde habbað englas to Gode þonne men
more affinity have angels to God than men
'angels have more affinity to God than people'

(Thorpe *Ælfric's Homilies* i.10)

b. *topic-pronominal subject-V*

æfter his gebede he ahof þæt cild up
after his prayer he lifted the child up
'after his prayer he lifted the child up' (ibid. ii.28)

c. *subject-V*

we schulon swiðe smealice ðissa ægðer underðencean
we must very narrowly these both consider
'we must consider both of these very carefully'

(*Pastoral Care* 48.23)

d. *Wh-V-subject*

Hwæt sægest þu yrþlincg?
what say you ploughman

'What do you say, ploughman?' (*Ælfric's Colloquy* 22)

When the sentence initial position is occupied by a topic, V precedes a full NP subject, whereas such inversion does not occur in the case of pronominal subjects, as shown in (34a, b). Van Kemenade (1987) argues that pronouns in OE are clitics which can freely adjoin to V. If this analysis is on the right track, we can maintain the assumption that finite verbs uniformly raise to *c* in OE. When the sentence initial position is occupied by a *wh*-phrase as in (34d), by contrast, V always precedes the subject irrespective of whether it is a full NP or a pronoun.[13]

The question we must ask is which pattern led children to set the V2 parameter. In inverted interrogatives, both pronominal and full NP subjects follow finite verbs, so that the V-subject word order is expected to have played a key role in fixing the V2 parameter (i.e. [+F] on *c*) for interrogative sentences.[14] On the other hand, the situation is a little more complicated in the case of declarative sentences in (34a-c). In his statistical analysis of V2 word order patterns, Gerritsen (1984) shows that in modern Dutch, German, and Norwegian, about 60 % of sentences begin with subjects in conversational styles. It is inferred from this figure that subject-initial sentences were dominant over topic-initial sentences in

colloquial OE as well. Moreover, given the fact that pronominal subjects precede V in topic-initial sentences, it seems quite natural to suppose that the majority of declarative sentences to which children were exposed exhibited the subject-V word order. For this reason, I will assume that children set the V2 parameter for declarative sentences on the basis of the subject-V word order and that the topic-V word order only served to confirm the result of the parameter setting.

One might argue that the subject-V word order cannot be a crucial factor in determining the V2 parameter because this order can be derived in non-V2 languages as well. Let us discuss this matter in some detail by taking into consideration the nature of [+F]. Recall our earlier discussion of finiteness being a concept that is prerequisite for tense and mood and guarantees the existence of predication at the time of utterances. This is to say that the notion of finiteness conceptually precedes specific values of tense: finiteness makes sure that a state of affairs expressed by V is on some point of the time axis and then tense determines specific values (e.g. present or past). Since [+finite] clauses are always [+tense], it is very difficult to sort out these notions in tensed clauses, but in the domain of non-finite clauses, the distinction manifests itself as observable syntactic effects. Following Stowell (1982), Martin (1992) classifies infinitival clauses in terms of the [±tense] feature of Infl. Here, [+tense] means that the non-finite Infl node has a time frame that is not simultaneous with the matrix clause and represents an

unrealized event (i.e. control infinitives), whereas [−tense] Infl does not have its own time frame so that it is interpreted as totally dependent on the matrix clause (i.e. ECM infinitives). Of these two types of infinitives, only those with [+tense] Infl license PRO in the subject position and allow VP-deletion to take place, as in (35):

(35) a. John wasn't sure he'd [$_{VP}$ win the race], but he tried [PRO to [$_{VP}$ *e*]]. ([+tense] infinitive)

b. *John considered [Mary to [$_{VP}$ be clever]] and Mike considered [Sally to [$_{VP}$ *e*]]. ([−tense] infinitive)

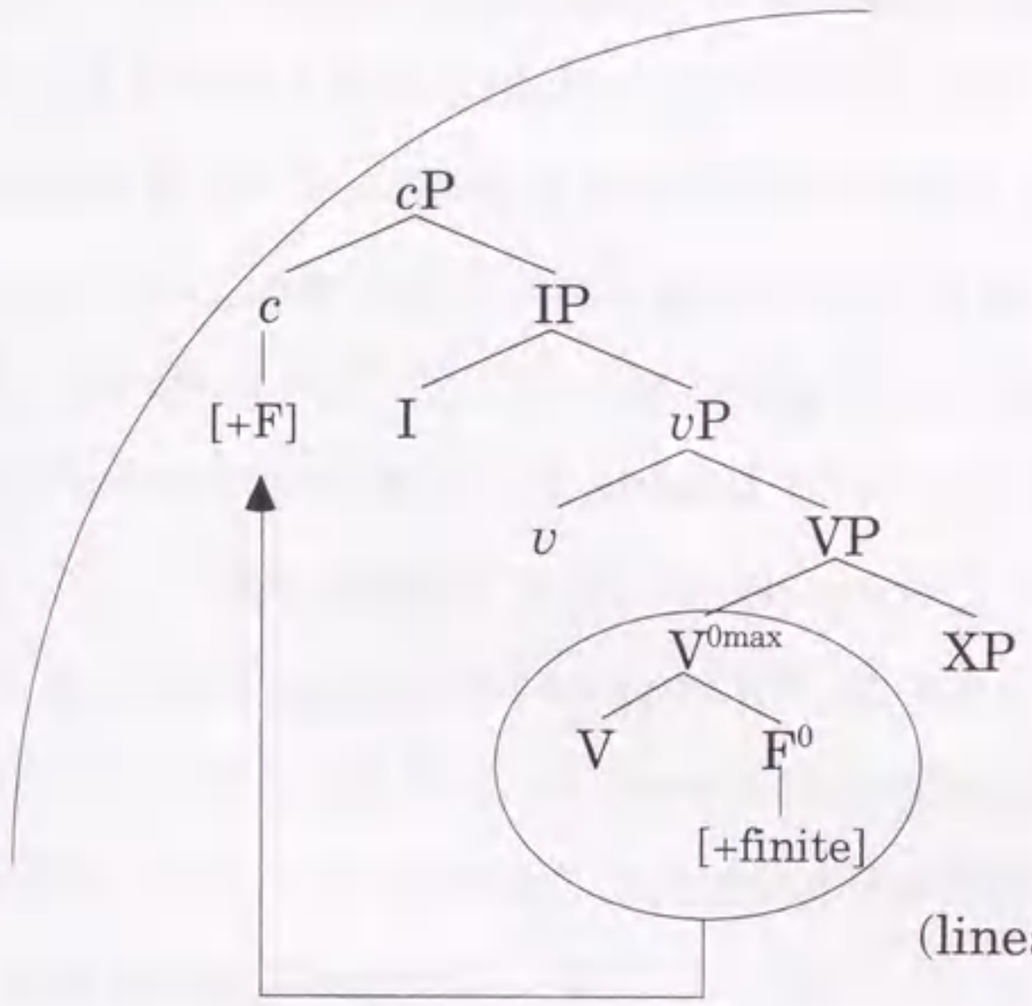
(Martin (1992:10-11))

This contrast shows that even [−finite] clauses can be specified as [+tense] and serves as evidence for the idea that the notion of finiteness conceptually precedes specific values of tense. Then, given this characterization, it follows that [+F], a feature that licenses finiteness of tensed clauses, takes scope over Infl at the interface level. This means that *c* is the *default* position of [+F] for the subject-V word order in declarative sentences. Therefore, once children are given the PLD that show the subject-V word order and a Degree III verbal inflectional system, they will automatically place [+F] on *c* as the default procedure.[15]

As we have seen in (16), OE and early ME had Degree III verbal morphology, in which both tense and agreement morphemes are found in past plural forms. Thus, children exposed to PLD of OE and early ME would judge that [+F] is on *c* in declarative sentences. Moreover, the V-subject word order of interrogative sentences

indicates that [+F] is unequivocally in *c*. The result is the following parameter setting for both declaratives and interrogatives:

(36) *OE ~ Early ME* (declaratives, interrogatives)



(linear order is irrelevant)

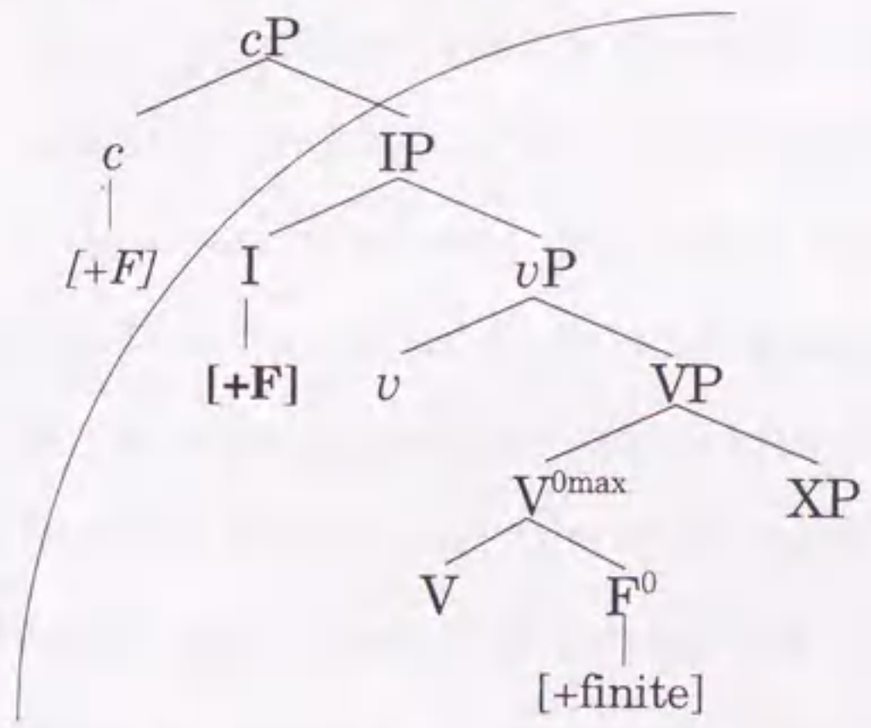
Since the Attraction Domain of Degree III morphology is *c*P, [+F] can identify the [+finite] feature. F^0 is a bound morpheme in the case of main verbs, so that when it moves to *c*, it carries along the whole V^{0max} (the intermediate steps of local head movement are omitted here) . This V-to-*c* movement yields V2 word order shown in (34). Alternatively, it is also possible that [+F] is checked off by a complementizer and V stays in situ, as argued in Chapter 3. This derives the *þæt*-S-O-V word order in embedded clauses.

The present hypothesis is confirmed by data from early child German. Boser et al. (1992) found that there is a strong correlation between verb positions and finiteness. Early children demonstrate

very productive V2 movement in German, but in all the examples where V fails to move to the V2 position and occurs in the finial position of main clauses, V bears the infinitival affix.[16] This suggests that it is not the category V but the finiteness feature which is sensitive to the V2 movement and that children know that whenever V is finite (and it is sufficiently rich), it does not appear at the original position and always moves to *c* in main clauses.

In the course of ME, V2 was being lost from the English grammar and almost died out around 1400 or in the first half of the 15th century. The point is that this is roughly the same period in which the verbal agreement morpheme *-en* for past plural forms was lost. In our terms, the loss *-en* means that English verbal morphology was recategorized as having the richness of Degree II, whose Attraction Domain is IP:

(37) *Late ME* (declaratives)



Since the [+finite] feature is now visible only within the domain of IP, [+F] cannot identify it outside this domain. Thus, children provided

with the PLD that show subject-V word order and Degree II verbal morphology would place [+F] on Infl. From the diachronic point of view, it appears as if [+F] shifts from *c* to Infl. It must be noted that this parameter resetting can be carried out without changing surface word order, given the assumption that only the dominant subject-V order plays a crucial role in fixing parametric values of declarative sentences. Thus, the shift of [+F] from *c* to Infl is consistent with the general tendency of language acquisition that the learner maximally preserves the surface pattern of PLD, as stated in (33).

Let us examine the process of diachronic feature shift in more detail. Strong operator features including [+wh] and [+F] fulfill dual functions at the same time: they induce overt movement of a *wh*-phrase or a finite verb in overt syntax and serve as scope markers at LF. As for the [+wh] feature (=Q feature), Chomsky (1995:289) observes that it "is plainly interpretable; therefore, like the ϕ-features of a nominal, it need not be checked—unless it is strong, in which case it must be checked before Spell-Out if a derivation is to be constructed." However, if a [+wh] feature is checked and deleted entirely, it cannot be construed as a scope marker at LF because a deleted feature is no longer visible at the interface level (ibid.: 280). Therefore, it is quite reasonable to suppose that a strong operator feature is a binary compound that consists of a strong sub-feature and an operator sub-feature. If this is on the right track, it can be said that only the strong sub-feature is checked off in overt syntax while the operator sub-feature remains intact and is properly

interpreted at LF.[17] The same analysis will apply to the strong [+F]. Since the notion of Attraction Domain is part of the definition of Attract F and has nothing to do with the scope marker interpretation, I propose that the decline of the morpheme *-en* caused only the shift of the strong sub-feature of [+F], whereas the operator sub-feature remained at *c* even after the mid-15th century. This gave rise to a split-distribution of [+F] in a single configuration. In (36), the strong sub-feature and the operator sub-feature are represented in bold type **[+F]** and in italic type *[+F]*, respectively. In overt syntax, a finite verb moves to Infl to check off **[+F]**, and its formal features further raise to *c* covertly to be properly interpreted (i.e. mapped into some point of the time axis) by means of *[+F]*:

(38) $[_{cP}$ *c* *[+F]* $[_{IP}$ Infl ~~[+F]~~ $[_{VP}$ V [+finite]]]]

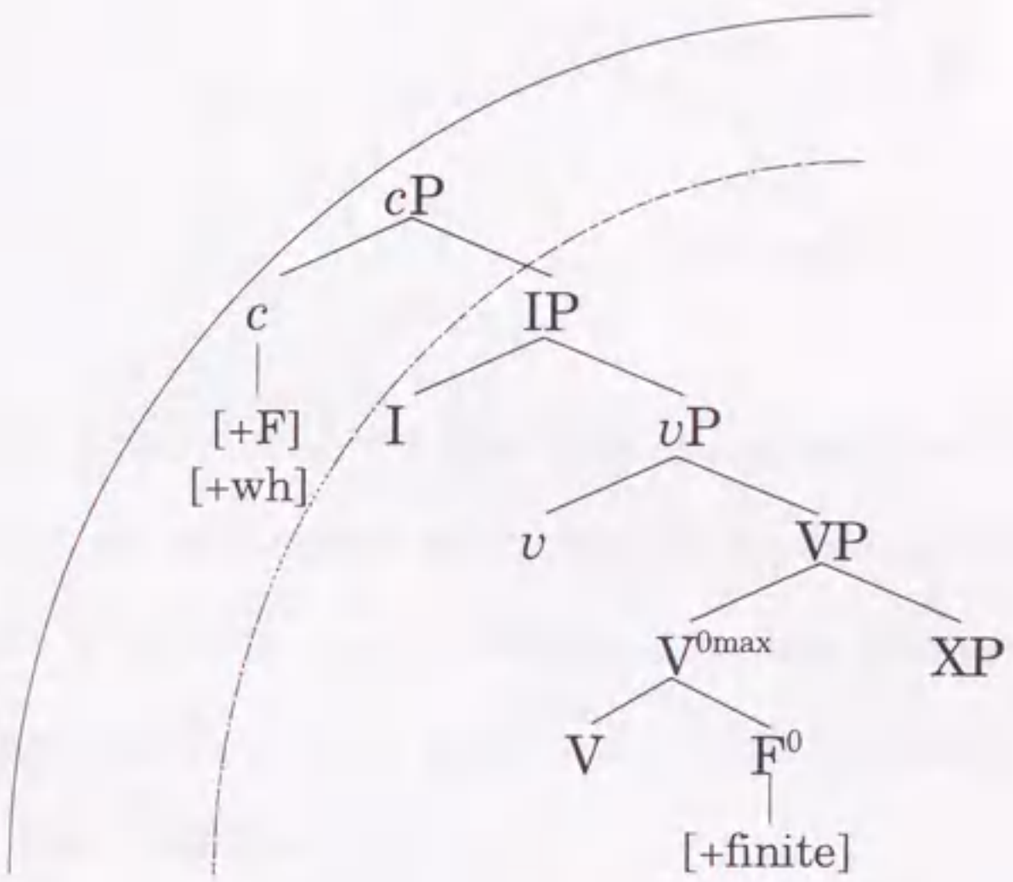
covertly *overtly*

A finite verb (or its formal features) finally raises to *c* regardless of its overt positions. This accords with the minimalist conceptual requirement of maximizing uniformity of LF outputs (cf. Chomsky (1995:362)).

We will now shift our attention to interrogative sentences. In inverted interrogatives in OE and early ME, the subject always follows the finite V. This word order can be derived only if [+F] is located in *c*; therefore, the shift of [+F] from *c* to Infl would inevitably change the surface word order. In view of the general tendency for children to preserve the surface word order of PLD, it is expected that children exposed to PLD of late ME with Degree II verbal

morphology continued to place [+F] on *c* in inverted interrogatives. Contrary to declarative sentences, fixing [+F] on *c* in late ME does not pose any problems in the case of inverted interrogatives.
Consider the following structure:

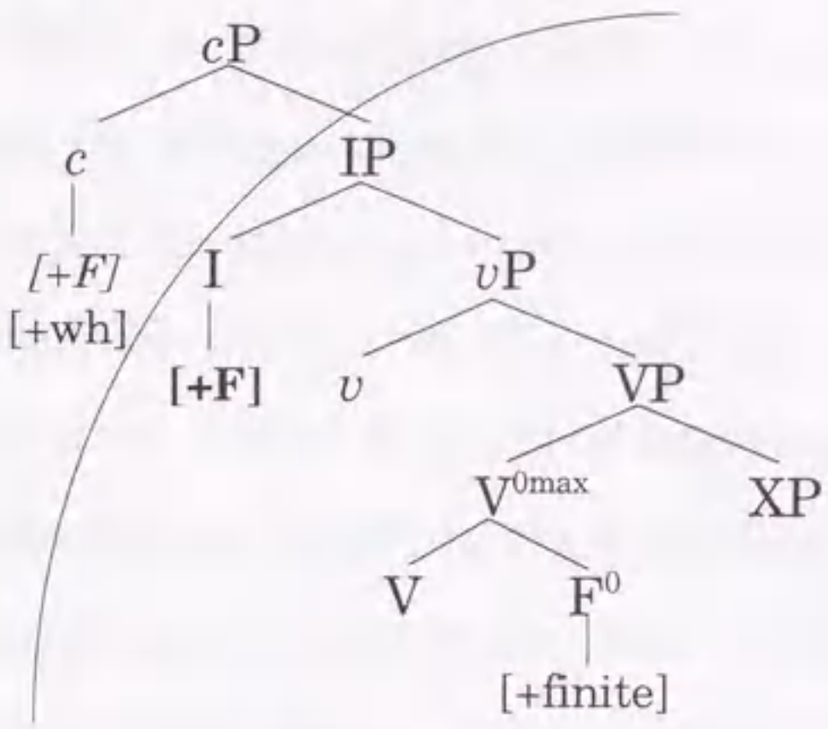
(39) *Late ME* (inverted interrogatives)



The point to note is that feature collaboration proposed in section 4.4.2 works here. The Attraction Domain of the Degree II verbal morphology in late ME is IP, but thanks to the [+wh] feature on *c*, the [+finite] feature behaves as if it had the strength of Degree III. Therefore, [+F] can identify the [+finite] feature from *c*.

On the other hand, in interrogative sentences where a *wh*-phrase itself is the subject occupying [Spec, *c*P], nothing prevents the shift of [+F] because there are no elements intervening between *c* and Infl. Diachronic feature shift thus can be implemented without changing the surface word order:

(40) *Late ME* (*wh*-subject interrogatives)



This shift brought about the subject/non-subject asymmetry in interrogative sentences with respect to the distribution of [+F]. This asymmetry will be overtly realized as the presence or absence of *do*-support in the latter half of the 16th century, to which we will turn in the next section.

Let us briefly look at diachronic feature shift in embedded clauses. As we have seen in Chapter 3, in OE and early ME, [+F] on embedded *c* could be checked off either by complementizer insertion or overt V raising (embedded V2 without complementizers), the latter option being restricted to complement clauses selected by certain bridge verbs. The Attraction Domain of the finite V is irrelevant in the case of complementizer insertion, since [+F] enters into a checking relation with the [+finite] feature of complementizers. Nevertheless, [+F] in embedded clauses shifted from *c* to lower head positions in much the same way as in main clauses. Lightfoot (1991) advances a hypothesis which maintains that the triggering

experience of parameter setting makes use of data only from unembedded domains (the degree-0 learnability). In accordance with this hypothesis, we assume that the parametric change in main clauses is reflected in embedded clauses as well. The only difference is that embedded clauses include C in addition to *c*. Thus, [+F] (more precisely, **[+F]**) first shifted to C at the beginning of the 14th century and then further shifted to Infl in the mid-15th century when the past plural morpheme *-en* was finally lost. As before, the operator sub-feature *[+F]* remained on *c*. This is supported by the fact that even in present-day English, the choice of complementizers is sensitive to finiteness of embedded clauses and only *that*, which has the [+finite] feature, is compatible with finite complement clauses; if the non-finite complementizer *for* is inserted, it will conflict with the *[+F]* specification of *c* and yield an interpretive mismatch.

To sum up the discussion in this section, we have seen that English verbal morphology was recategorized as the class of Degree II morphology in the first half of the 15th century due to the decline of the past plural morpheme *-en*, and that this decline caused the shift of [+F] from *c* to Infl, except for inverted interrogatives where the subject intervenes between these two head positions. In inverted interrogatives, [+F] remained on *c* and continued to induce long-distance V-raising. Therefore, the proposed mechanism naturally captures the insight that subject-auxiliary inversion in present-day English is literally the residue of V2 that dates back to OE.

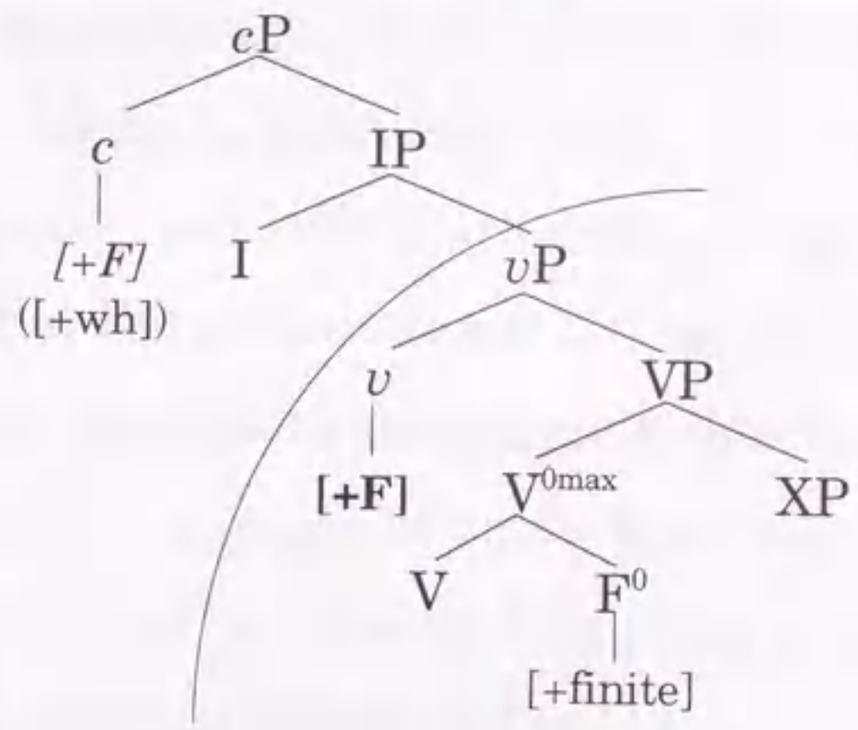
It is also conceptually important that we evade here the logical

problem of language change raised in Chapter 1, by assuming that the parametric change was triggered by the erosion of verbal morphology. If the topic-V-subject word order was the determining factor for selecting the V2 option and the decreased amount of this order forced children to reset the parameter to the negative value as many authors claim, then the first decrease of topic-initial sentences will remain unexplained. Instead, we have argued that children could set the V2 parameter for declarative sentences on the basis of the dominant subject-V word order since *c* is the default position of [+F]. Thus, as long as English retained the Degree III verbal morphology, children placed [+F] on *c*, but once that rich morphology was lost because of language contact and other extra-linguistic reasons, they began to place it on Infl without changing the surface word order of PLD. Under the present system, therefore, the decrease of the topic-initial sentences is considered to be the result, not the cause, of the loss of V2.

4.5.2. The Loss of V-to-I

Later in the mid-16th century, the verbal agreement morpheme *-est* for past second singular was lost (cf. Pyles (1964:205), Strang (1970:203)). As we have seen in section 4.3.2, this decline of verbal morphology coincides with the loss of V-to-I movement. In our terms, the demise of *-est* means that English verbal morphology was recategorized as having the richness of Degree I, whose Attraction Domain is *v*P as shown in (41):

(41) *Early ModE* (affirmative declaratives, *wh*-subject interrogatives)



Since the [+finite] feature is visible only within this domain, children exposed to PLD with Degree I verbal morphology would place **[+F]** on *v* in affirmative declaratives and *wh*-subject interrogatives, where nothing intervenes between Infl and *v* and thus parameter resetting was properly carried out without changing surface word order patterns.

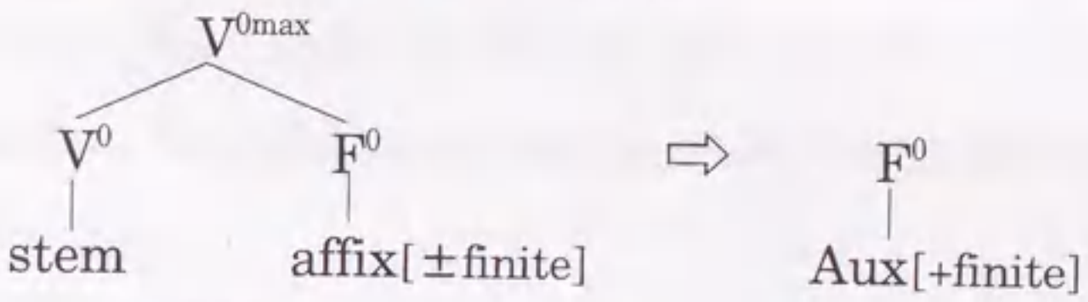
However, it is important to note that the shift of **[+F]** did not occur in all types of sentences. **[+F]** remained in Infl or *c* when there was no need to place it on *v*, or when some intervening element blocked the diachronic feature shift. The former case was sentences involving auxiliaries. It is generally assumed in the literature (Lightfoot (1979, 1991), Nakano (1985, 1994), Roberts (1993) *inter alia*) that modal verbs were reanalyzed at the beginning of the 16th century as a distinct group of verbs that occur only in Infl. Though we will not go into details of this reanalysis, it was most plausibly

induced by the loss of infinitival affix *-an* (*-en*, *-on*), which enabled children to interpret modal verbs not as IP-taking control verbs but as *v*P-taking Infl elements, as the authors cited above point out:

(42) Modal $[_{IP}$ PRO Infl *v*P $]$ ⇨ $[_{IP}$ Modal $[_{vP}$ *v* VP $]]$

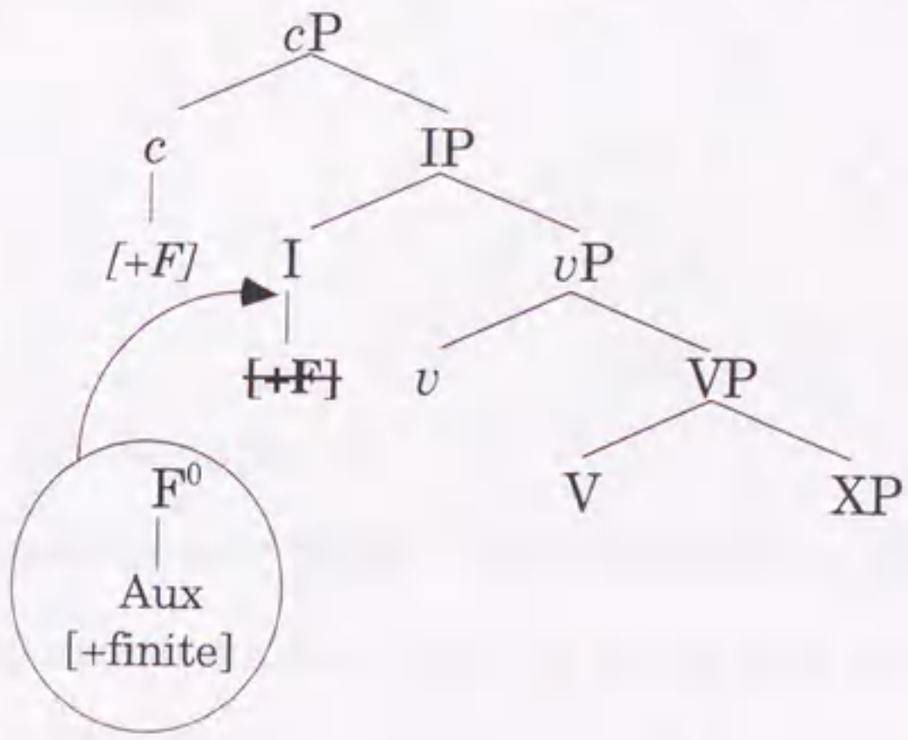
We have argued in section 4.3.1 that stems and affixes of main verbs separately enter the computational system whereas auxiliaries are always thrown into phrase structures in their full forms. Thus, it is expected that the reanalysis of modal verbs also affected their internal structure. They came to be directly dominated by F^0 bearing the [+finite] feature, as illustrated in (43):

(43)



Once modals were established as a distinct category from main verbs, they came to be directly merged with Infl.[18] Then, **[+F]** on Infl could be checked off against the [+finite] feature of auxiliaries:

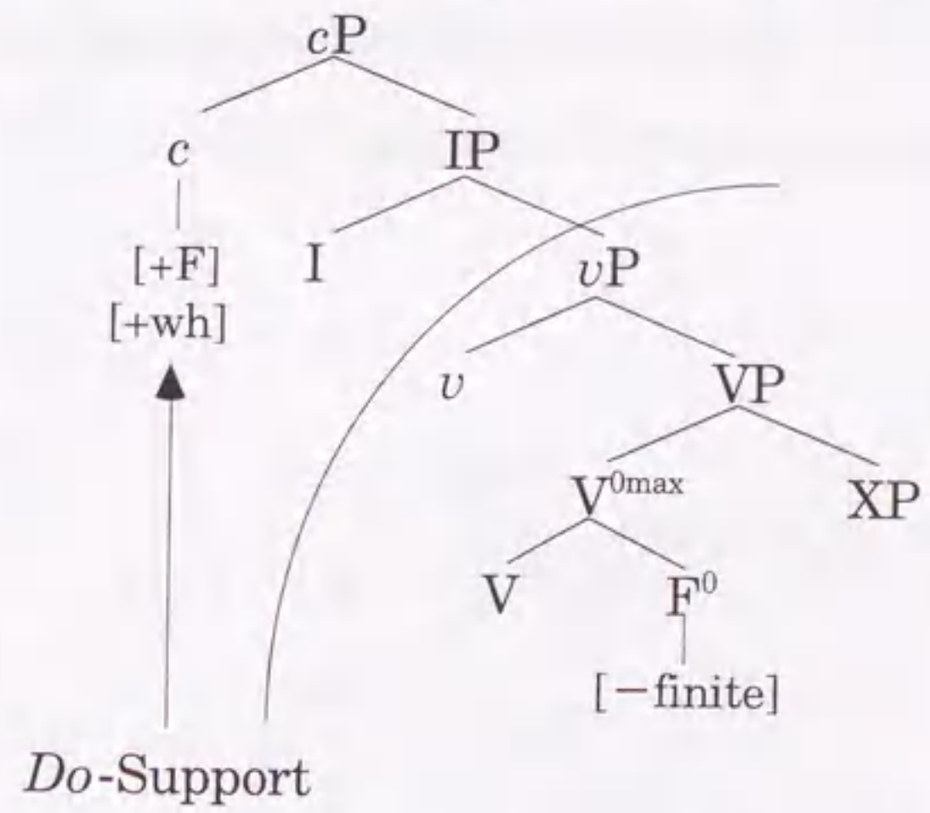
(44) *Early ModE* (sentences involving auxiliaries)



Note that the Attraction Domain is irrelevant in the case of insertion of F^0, since it is part of the definition of Attract F, but not Merge. As a result, even after English verbal morphology eroded to be recategorized as the class of Degree I, **[+F]** remained in Infl in sentences involving auxiliaries.

Let us turn to cases where diachronic feature shift did not take place because of certain intervening elements. Recall here that in late ME, children failed to place [+F] on Infl when the subject occupied [Spec, IP] in inverted interrogatives. The same is true in the mid-16th century as well. Because parameter setting that would change the surface word order of PLD is not favored in view of the general tendency of language acquisition, [+F] continued to occupy *c* in early Modern English:

(45) *Early ModE* (inverted interrogatives)

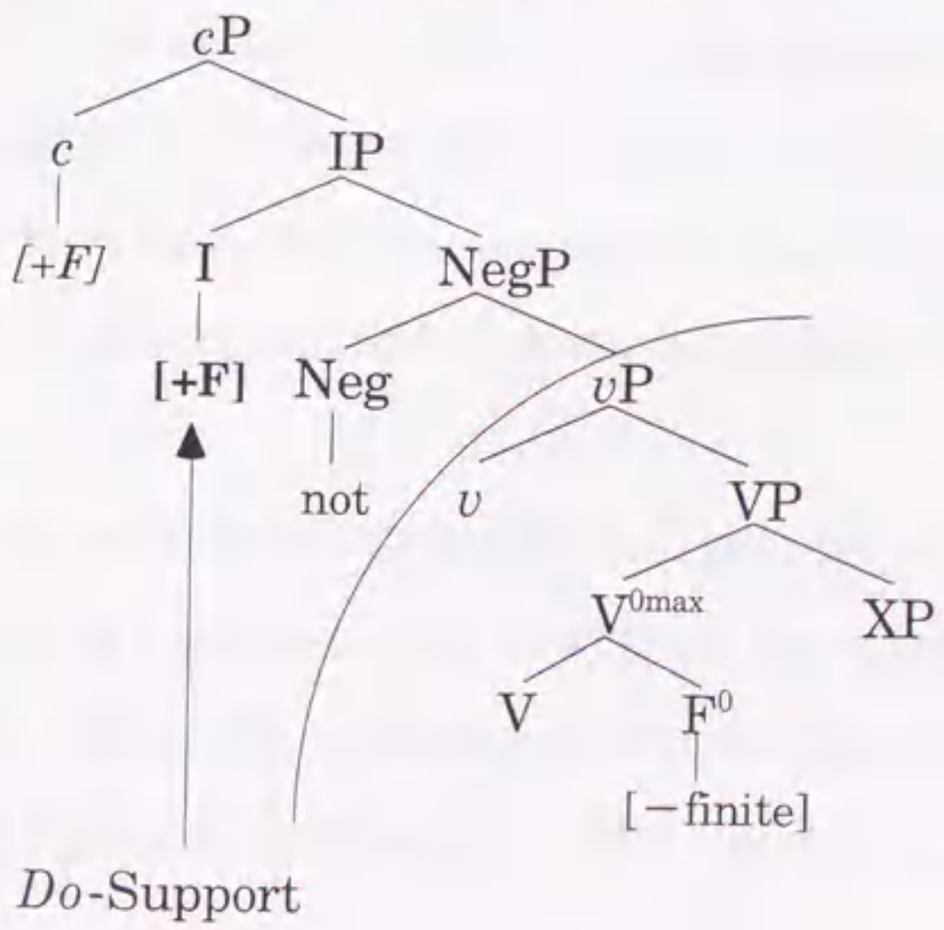


The situation in early ModE, however, differs from that in late ME in that feature collaboration does not work well any longer. Since the Attraction Domain of the Degree I verbal morphology in early ModE

is *v*P, the supporting feature [+wh] cannot extend the Attraction Domain to *c*P because of the presence of the maximal projection IP (see the definition in (29)). Features of the verbal affix are therefore invisible to [+F]. If [+F] remains unchecked, the derivation is bound to crash, but there was a way to save the structure in (45). Along with modal verbs, *do* also underwent the process of reanalysis at the beginning of the 16th century and became a member of auxiliaries (cf. Ellegård (1953) and references cited above). Thus, in inverted interrogatives without modals nor *have/be*, the dummy *do* came to be employed as the last resort to check off [+F].

This analysis can be extended to *do*-support in negative sentences involving *not*. We assume that *not* is generated as the head of NegP, following Pollock (1989) and Chomsky (1991). This intervening element blocks the shift of **[+F]** from Infl to *v* for the same reason as in inverted interrogatives:

(46) *Early ModE* (negative sentences involving *not*)



[+F] cannot identify features of the verbal affix from outside its Attraction Domain, so that if a sentence does not employ modals nor *have/be*, **[+F]** must be checked off by inserting *do* as a dummy element.

As for emphatic sentences, let us assume that phonological features responsible for emphatic stress are carried by a functional category Emp and that these phonological features are assigned to a verbal element that occupies Infl in overt syntax:

(47) $[_{cP}\ c\ [_{IP}\ \text{Infl}\ [_{EmpP}\ \text{Emp}\ [_{vP}\ v\ [_{VP}\ \text{V}\ ]]]]]$

stress

If V only raises to *v* and phonological features of Emp stand alone, the derivation will crash at the PF side. Thus, to guarantee that Infl was filled with some lexical element, children continued to place **[+F]** on Infl even after mid-16th century. When no other auxiliaries are employed, *do* is inserted to check off **[+F]** and thereby carries the emphatic stress. Summarizing the discussion so far, we have seen that **[+F]** shifted to *v* in affirmative declaratives and *wh*-subject interrogatives but continued to occupy *c* or Infl in the other types of constructions (inverted interrogatives, negative sentences, emphatic sentences, and sentences involving auxiliaries) after the mid-16th century.

Some additional explanation is needed with respect to examples which are apparently unexpected from the analysis above. First, let us consider *do* in affirmative declaratives without emphasis. As the following figure from Ellegård (1953) indicates, *do*-support was

optionally used in affirmative declaratives during the 16th century, but after the 17th century, *do* rapidly declined in affirmative declaratives, as opposed to the other types of sentences:

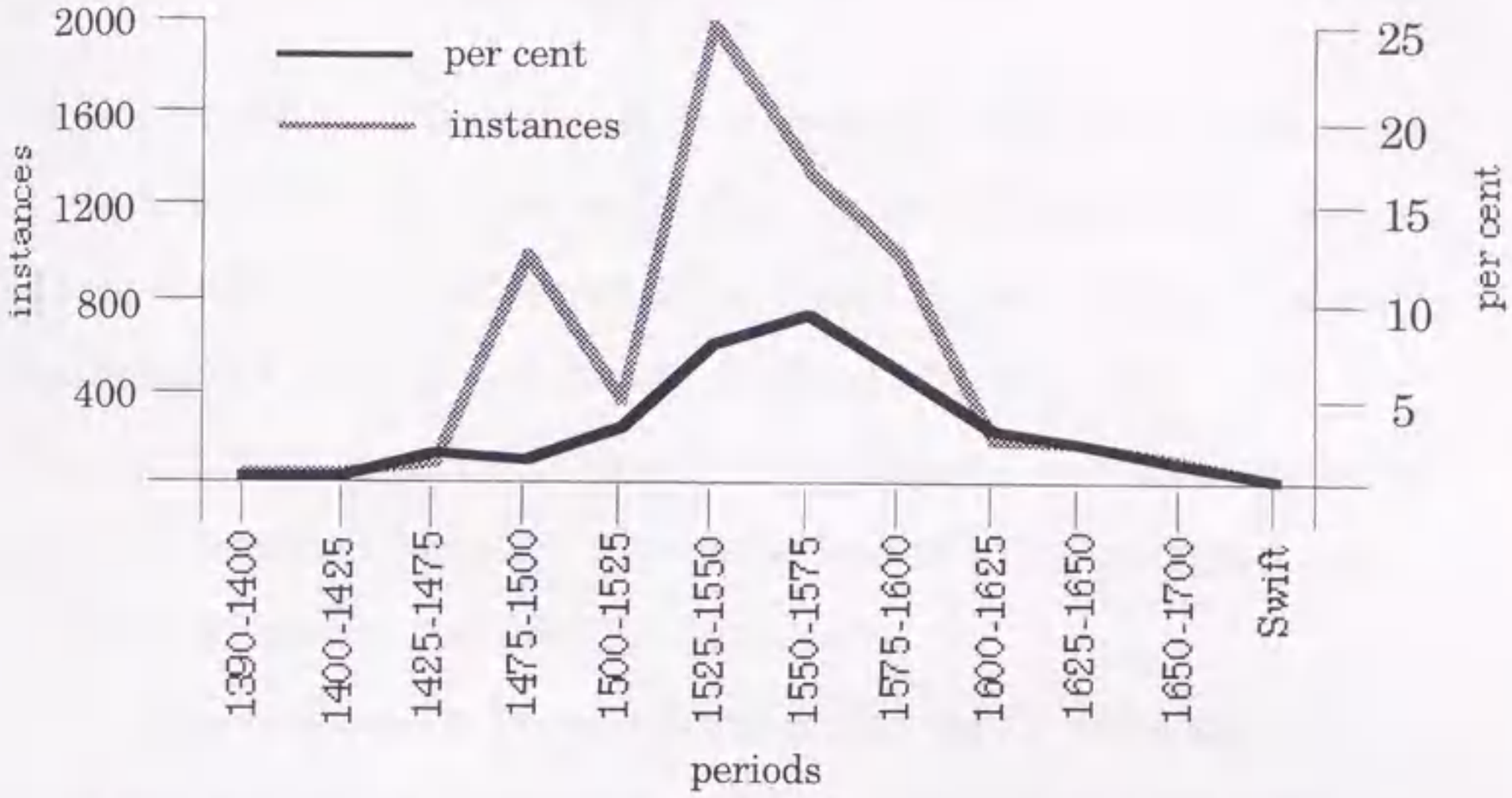


Figure: Number and per cent of do *in affirmative declaratives*

(cf. Ellegård (1953:161))

This phenomenon can be adequately accounted for within the present framework. If we follow Lightfoot (1979, 1991) and Roberts (1993) on the reanalysis of *do* and Roberts (1993) and Vikner (1997) on the loss of V-to-I movement, the former occurred at the beginning of the 16th century, while the latter change took place around 1575. It is thus predicted that from 1500 to 1575, both *do*-support and overt V-raising were available as free variations to check off [+F] in all types of sentences, and *do*-support in affirmative declaratives indeed is most frequently observed in this period, as illustrated above. The two derivations are represented as follows:

(48) a. did [$_{IP}$ Mary$_i$ Infl [$_{vP}$ t_i see John]] (*do*-insertion)
~~[+F]~~

b. [$_{IP}$ Mary$_i$ Infl [$_{vP}$ t_i saw John]] (V-raising)
~~[+F]~~

Notice that since these two derivations have different numerations, they do not compete with each other. After [+F] shifted to *v* around 1575, it was no longer necessary to insert *do* into Infl in affirmative declaratives, so that *do*-support declined. On the other hand, in constructions where [+F] stayed in the higher functional heads, *do*-support became obligatory due to the loss of V-to-I movement and thus its frequency remarkably increased.

A second case to be considered is the *not*+V word order in negative sentences. According to Ukaji (1992), this minor but significant pattern began to appear at the beginning of the 15th century, flourished in the times of Shakespeare, and finally declined in the mid-18th century. Examples from each period are given below:

(49) a. I *Not holde* agaynes luste al vttirly.
I not hold against wishes all utterly
(c 1412 *The Regement of Princes*, Ukaji (1992: 454))

b. For who *not needs* shall never lack a friend,
(Shakespeare *Hamlet* III.ii.207, ibid.)

c. They . . . possessed the island, but *not enjoyed* it.
(1740 *Life Drake*; *Wks*. IV.419, ibid.)

It was argued above that finite verbs must always precede *not* for the purpose of checking **[+F]** which remains in Infl even after the mid-

16th century in negative sentences. The *not*+V word order above constitutes an apparent counterexample to this explanation.

To settle this problem, let us focus on the status of adverbs in diachronic feature shift. While V raised to Infl in late ME, VP adverbs like *often* followed finite verbs, but later in early ModE, when **[+F]** had shifted to v, these adverbs came to precede them (see (31) and (32) in Chapter 3). Since adverbs are optional elements, it seems entirely fair to say that they had no blocking effect on diachronic feature shift and children could shift the position of **[+F]** in affirmative declaratives ignoring them. It is interesting to note that negative markers *never*, *scarcely*, and *seldom* also precede finite verbs and do not require *do*-support in present-day English:

(50) a. The princess [$_{vP}$ never smiled again].

b. I [$_{vP}$ scarcely understand him].

c. Mrs. Green [$_{vP}$ seldom goes out].

Pollock (1989:405) argues that these negative markers occupy VP-initial adverbial positions. If this is the case, we can deal with these examples on a par with affirmative declaratives involving VP adverbs, assuming that **[+F]** is located in v. Then, I would suppose that the *not*+V word order in (49) has the same structure as (50), with *not* in the vP-adjoined position and **[+F]** being located in v. In his analysis of the historical development of *not*, Ishikawa (1995:209) states that *not* "began to be reanalyzed as an X^0 category and generated in Neg^0 in the 16th century, although XP *not* also existed." If this XP *not* had the adverbial status like *never*, the *not*+V word

order is naturally derived.[19] The fact that Infl elements like modals never appeared to the right of *not* (Ukaji (1992:456)) suggests that XP *not* is adjoined to *v*P. Afterwards, when *not* was completely reanalyzed as X^0 and XP *not* died out, this word order also declined.

4.6. *Do*-Support in Present-Day English

So far, we have seen how the seemingly anomalous distribution of [+F] in present-day English illustrated in (5), repeated here as (51), has been brought about as a result of the diachronic feature shift driven by the loss of verbal morphology:[20]

(51) a. In affirmative declaratives, matrix *wh*-subject interrogatives, and embedded interrogatives in general, [+F] is located no higher than *v*.

b. In negative sentences, emphatic sentences, and sentences including auxiliaries, [+F] is located in Infl.

c. In matrix inverted interrogatives, [+F] is located in *c*.

With this distribution in mind, let us verify how the derivations of negative sentences, emphatic sentences and inverted interrogatives proceed in present-day English.

The basic clause structures of these types of sentences are delineated as follows:

(52) a. $[_{IP}$ Subj Infl$_{[+F]}$ $[_{NegP}$ not $[_{vP}$ *v* $[_{VP}$ V[−finite]]]]]
(negative sentences)

b. $[_{IP}$ Subj Infl$_{[+F]}$ $[_{EmpP}$ Emp $[_{vP}$ *v* $[_{VP}$ V[−finite]]]]]
(emphatic sentences)

c. $[_{cP}$ Wh c $[_{IP}$ Subj Infl $[_{vP}$ v $[_{VP}$ V[−finite]]]]]
[+F]
(inverted interrogatives)

[+F] on Infl and c must be checked off against a [+finite] feature. However, [+F] cannot attract a possible candidate V in the phrase structures because the Attraction Domain of verbal affixes in present-day English is vP, and they are invisible to [+F] in higher head positions. Thus, V is always [−finite]. The only way to check off [+F] is to directly insert some lexical element with the [+finite] feature into Infl or c by Merge. If a modal is included in the numeration, it can serve as the checker; otherwise, the dummy *do* is employed. Note also that Infl has tense and ϕ-features in addition to [+F]. These features are checked off against corresponding tense and ϕ-features of a modal or *do*. In (52c), tense and ϕ-features of Infl get checked by raising to c covertly. Since ϕ-features of verbs are uninterpretable (Chomsky (1995:278)), a modal or *do* inserted into c can serve as an attractor that raises formal features of Infl.

When the aspectual auxiliaries *have/be* are involved, they are generated as the head of FP:

(53) a. $[_{IP}$ Subj Infl $[_{NegP}$ not $[_{FP}$ have/be[+finite] $[_{vP}$ v $[_{VP}$ V]]]]]
[+F]
(negative sentences)

b. $[_{IP}$ Subj Infl $[_{EmpP}$ Emp $[_{FP}$ have/be[+finite] $[_{vP}$ v $[_{VP}$ V]]]]]
[+F]
(emphatic sentences)

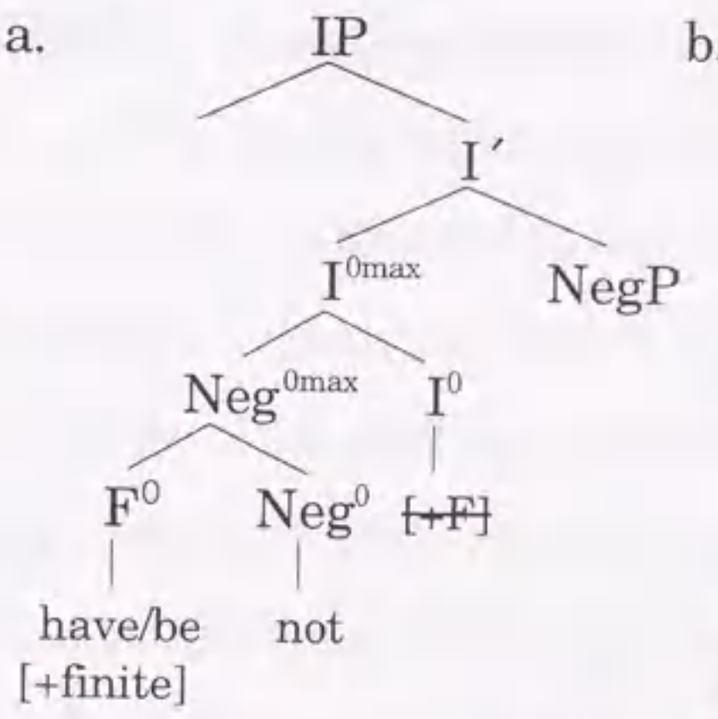
c. $[_{cP}$ Wh c $[_{IP}$ Subj Infl $[_{FP}$ have/be[+finite] $[_{vP}$ v $[_{VP}$ V]]]]]
[+F]
(inverted interrogatives)

As we have seen in section 4.3.1, *have/be* have no internal branching

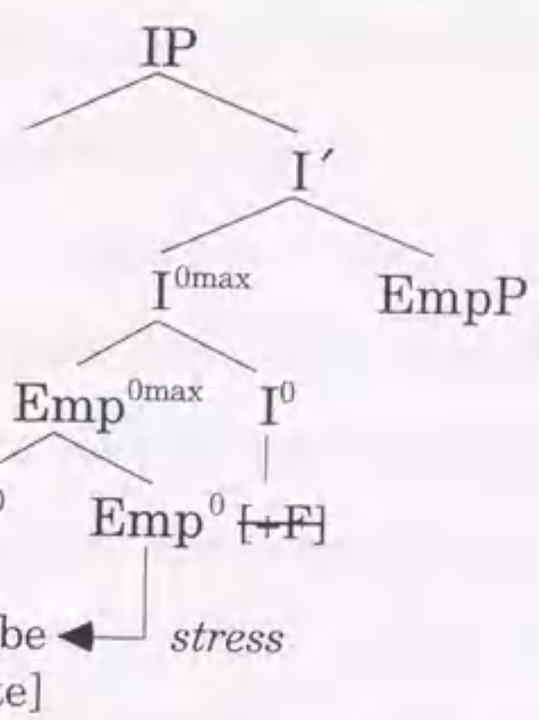
and bear the [+finite] feature by themselves; therefore, contrary to main verbs, F^0 is non-affixal in (53). From the definition in (26), the Attraction Domain of a non-affixal X^0 is the root. Thus, even if a modal is not included in the numeration, [+F] can identify the [+finite] feature of F^0 and attract *have/be* to its checking domain. Tense and ϕ-features of Infl are also checked off against corresponding features of *have/be*.

Note that NegP and EmpP intervene between IP and FP in negative sentences and emphatic sentences, respectively. Movement of *have/be* across Neg or Emp will lead to the violation of the shortest move requirement of Attract F stated in (25bi). To meet this requirement, *have/be* first raise to Neg or Emp, and then the amalgamates further raise to Infl. The structures of resulting complex heads look like (54):

(54) a.



b.



Suppose here that features contained in an X^{0max} that are adjoined to the target K are in the checking domain of K.[21] In (54), the [+finite] features contained in Neg^{0max} and Emp^{0max} adjoined to Infl are in the checking domain of Infl. Therefore, [+F] is properly checked off

against the [+finite] features in these configurations. In addition, since F^0 is adjoined to the left of Neg^0 in (54a), the correct word order is properly derived. In emphatic sentences, stress is assigned to *have/be* within Emp^{0max}, as illustrated in (54b).

A problem is that the option of inserting *do* is not available when *have/be* are involved in the phrase structure. The following sentences are totally ungrammatical:

(55) a. *Mary does not have seen John.

b. *Mary DOES have seen John.

c. *Does Mary have seen John?

If the numeration includes the dummy *do* and its [+finite] feature checks off [+F], derivations of these examples would incorrectly converge. Here, the parallelism between *do* and the expletive *there* mentioned in section 4.2 above provides a key to resolve this problem. Chomsky (1991) argues that *there* is not a legitimate LF object and if it remains at LF it causes a violation of the principle of Full Interpretation (FI). To avoid this violation, an associate with semantic contents covertly adjoins to *there*. Since the dummy *do* also lacks semantic contents and can be regarded as the X^0 counterpart of the expletive *there*, the same analysis will apply to the present case. Then, I would propose the following condition:

(56) In covert syntax, the dummy *do* is replaced by a bare infinitive with semantic contents to satisfy FI.

Here, I use the term "semantic contents" to cover the notions of argument structure and state of affairs designated by the infinitive.

If *have/be* appear as bare infinitives as in (55), they will replace *do* in covert syntax. However, these elements specify only aspects (or passive voice) and do not have argument structures nor represent state of affairs. Therefore, their adjunction to the dummy *do* will lead to a violation of FI, resulting in the ungrammaticality.[22]

So far, little has been said about occurrences of *do* in imperative sentences. As in declarative sentences, *do*-support is employed in negative imperatives and emphatic imperatives:

(57) a. *affirmative imperatives*

Tell me the truth.

b. *negative imperatives*

Don't forget these words.

c. *emphatic imperatives*

Do sit down!

As for the distribution of [+F], it seems reasonable to suppose that it occupies the same positions as in corresponding declarative sentences; thus, (i) in affirmative imperatives, [+F] is located in v and (ii) in negative imperatives and emphatic imperatives, [+F] is located in Infl. Furthermore, following Lasnik (1981, 1992), we assume that Infl in imperative sentences lacks tense features and carries an imperative affix -IMP as an alternative. The structures of (57a-c) are represented in (58a-c):

(58) a. $[_{IP}$ *pro* Infl $[_{vP}$ v $[_{VP}$ V[+finite]]]]

-IMP ~~[+F]~~

(affirmative imperatives)

b. $[_{IP}$ *pro* Infl $[_{NegP}$ not $[_{v\,P}$ *v* $[_{VP}$ V[−finite]]]]]
~~[+F]~~
-IMP (negative imperatives)
Do-Support

c. $[_{IP}$ *pro* Infl $[_{EmpP}$ Emp $[_{v\,P}$ *v* $[_{VP}$ V[−finite]]]]]
~~[+F]~~
-IMP (emphatic imperatives)
Do-Support

Due to its affixal nature, -IMP must be supported by some lexical item in overt syntax. In affirmative imperatives, finite V first raises to *v* to check off [+F], and then overtly raises to Infl to support the affix -IMP. In negative and emphatic imperatives, on the other hand, V-raising cannot be resorted since V is invisible to [+F], so that the dummy *do* is directly inserted into Infl and checks off [+F] and supports -IMP at a time.

The central problem of *do*-support in imperative sentences is its cooccurrence with auxiliaries *have/be*. The ungrammaticality of (59b, d) shows that contrary to declarative sentences, the dummy *do* must occur in negative imperatives even when *have* or *be* appears:[23]

(59) a. Don't be frightened by that noise.
b. *Be not frightened by that noise.
c. Don't have finished your homework when I come back!
d. *Have not finished your homework when I come back!
(cf. Quirk et al. (1985:134), Arimoto (1994:248))

The same holds true of emphatic imperatives as well. Stress is always assigned to the dummy *do*, even if *have* or *be* appears as the leftmost verbal element:

(60) a. DO be seated.

b. *BE seated.

c. DO have finished your homework when I come back!

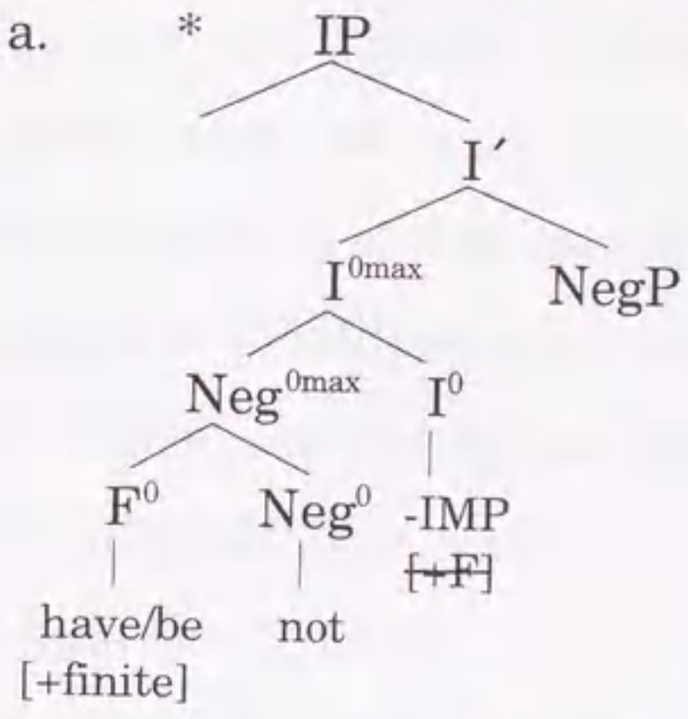
d. *HAVE finished your homework when I come back!

(cf. Quirk et al. (1985:833), Arimoto (1994:249))

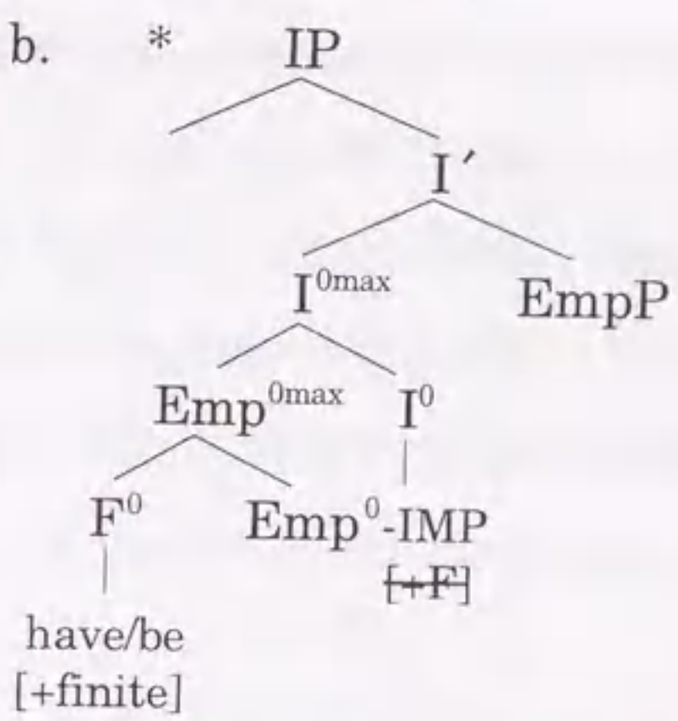
It must be made clear why *have/be* raising is blocked in favor of *do*-support only in imperatives.

I suggest that the unavailability of *have/be* raising in the examples above can be attributed to the affixal property of -IMP. If F^0 dominating *have/be* moves successive-cyclically to Infl as in declarative sentences, the following structures will result:

(61) a.



b.



As before, [+F] can be properly checked off against the [+finite] feature. What is problematic with these configurations is the licensing of -IMP. Since -IMP is the imperative affix, it must be attached to a verbal stem. However, in (61a, b), F^0 is dominated by Neg^{0max} or Emp^{0max}, and -IMP is attached from outside these complex heads. Thus, it cannot have a direct connection with the stem *have/be*, which causes the derivations to crash. The only way to

support -IMP, then, is to insert *do* into Infl directly; hence the obligatory occurrences of *do* in the examples above.[24, 25]

Finally, let us briefly consider *do*-support in VP-ellipsis. As is well-known, VP-ellipsis in English is possible only when Infl is filled with some auxiliary. If no modals nor *have/be* are employed, the dummy *do* is obligatorily inserted, as shown in (62):

(62) a. Mary did not leave but John did [$_{VP}$ *e*].

b. Mary left but John did not [$_{VP}$ *e*].

When an elided VP occurs in the complement of Neg as in (62b), *do*-support can be accounted for on a par with ordinary negative sentences, since [+F] located in Infl must be checked off by an auxiliary in any event. More problematic are cases like (60a), where an elided VP is contained in an affirmative declarative clause in which [+F] is located in *v*. One possible approach that comes to mind is to consider that VP-ellipsis affects only the lower VP of the Larsonian VP shell and *did* is inserted to check off [+F] in the upper *v* in (62a). This possibility, however, is immediately rejected in the face of the following example:

(63) A: Does Bob usually walk to work?

B: No, but his sister *does*. [= usually walks to work]

(Quirk, et al. (1985:907))

Does in the second sentence corresponds to *usually walks to work* in the first sentence. Given the widely accepted view that VP adverbs like *usually* mark the leftmost edge of *v*P, it is clear that VP-ellipsis affects the whole *v*P together with [+F] in (63). Nevertheless, *do*-

support is obligatory in the second sentence.

Thus, we are led to conclude that *do*-support in (62a) is not triggered by strong [+F]. Here, it is important to recall that Infl has tense and ϕ-features apart from [+F]. In ordinary affirmative declaratives, these features are checked off through covert feature raising of main verbs, but in the case of VP-ellipsis, in which phonological and syntactic features are made invisible (cf. section 4.3.1), tense and ϕ-features of Infl will be left unchecked and cause the derivation to crash if no auxiliary fills the position of Infl. Therefore, when no modals nor *have/be* are included in the numeration, the dummy *do* is employed to save the derivation. Since tense and ϕ-features of Infl are weak, checking actually takes place in covert syntax, but insertion of *do* must be overt. This is because covert insertion of a lexical item is barred if it has phonological features; unless it is not introduced into the phrase structure before Spell-Out, phonological features cannot enter the PF side and yield a violation of FI (cf. Chomsky (1995:231)).

We have seen that *do*-support in inverted interrogatives and negative/emphatic sentences was established around 1575. In the case of VP-ellipsis, by contrast, *do*-support was already employed as early as in OE. Some examples are given below:

(64) a. & sæden ðat micel þing sculde cumen herefter: sua
and said that great thing should come hereafter as
dide, for þat ilc gær warth þe king ded
did for that very year became the king dead
'and said that a major event should follow, which it did,

for that very year the kind died'

(*Peterb.Chon.* 1135.5, Denison (1993:272))

b. Hwi nolde god him forgyldan his bearn
why would-not God him compensate-Inf his children
be twyfealdum, swa swa he *dyde* his æhta?
by twice as as he did his possessions
'Why would not God compensate him for his children twice over, just as he did for his possessions?

(Thorpe *Ælfric's Homilies* ii.267.261,Warner(1993:118))

In OE, [+F] is located in *c* and tense and φ-features are independently carried by Infl. In ordinary sentences, these features are properly checked off by main verbs, because verbal inflections are still very rich and [+F] can attract main verbs. However, when VP-ellipsis is applied, these features will remain unchecked unless a modal-like element is inserted into Infl and further raises to *c*. I suspect that the use of *do* in cases like (64) prepared the ground for the spread of the dummy *do* to the other types of sentences in the later period, but a full investigation of the origin of *do* as a dummy element lies outside the scope of this thesis.

4.7. A Residual Problem: V2 in Mainland Scandinavian

In this chapter, it has been argued that the richness of verbal morphology and the possibility of long-distance verb movement correlate with each other and shown that both of them became out of use hand in hand in the history of English. However, as we have touched on in section 4.3.2, rich verbal morphology is not a necessary

condition on V-to-*c* movement. There are in fact languages with flat morphology that exhibit V-to-*c* movement: namely, mainland Scandinavian languages (Swedish, Norwegian, Danish). These languages have no agreement morphology in both present and past tenses, as illustrated below:

(65) *Verbal Inflections in Mainland Scandinavian*

	Swedish		*Norwegian*		*Danish*	
	Present	Past	Present	Past	Present	Past
Sg. 1	-ar	-ade	-er	-et	-er	-ede
2	-ar	-ade	-er	-et	-er	-ede
3	-ar	-ade	-er	-et	-er	-ede
Pl. 1	-ar	-ade	-er	-et	-er	-ede
2	-ar	-ade	-er	-et	-er	-ede
3	-ar	-ade	-er	-et	-er	-ede

In the present framework, verbal morphologies of these languages are grouped into the class of Degree I, so that it is predicted that finite verbs cannot move as far as *c*. This prediction, however, does not borne out. V2 word order is observed in these languages:

(66) a. *Swedish*

I Sverige har han många vänner.
In Sweden has he many friends

'He has many friends in Sweden.'

b. *Norwegian*

Altså kommer han ikke i morgen.
so come he not tomorrow

'So he will not come tomorrow.'

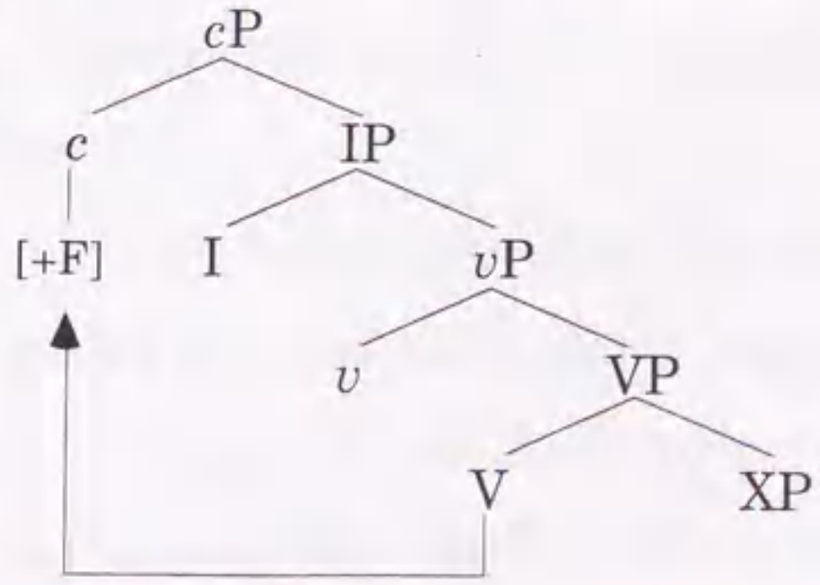
c. *Danish*

Denne bog gav mor hende i går.
this book gave mother her yesterday
'Mother gave her this book yesterday.'

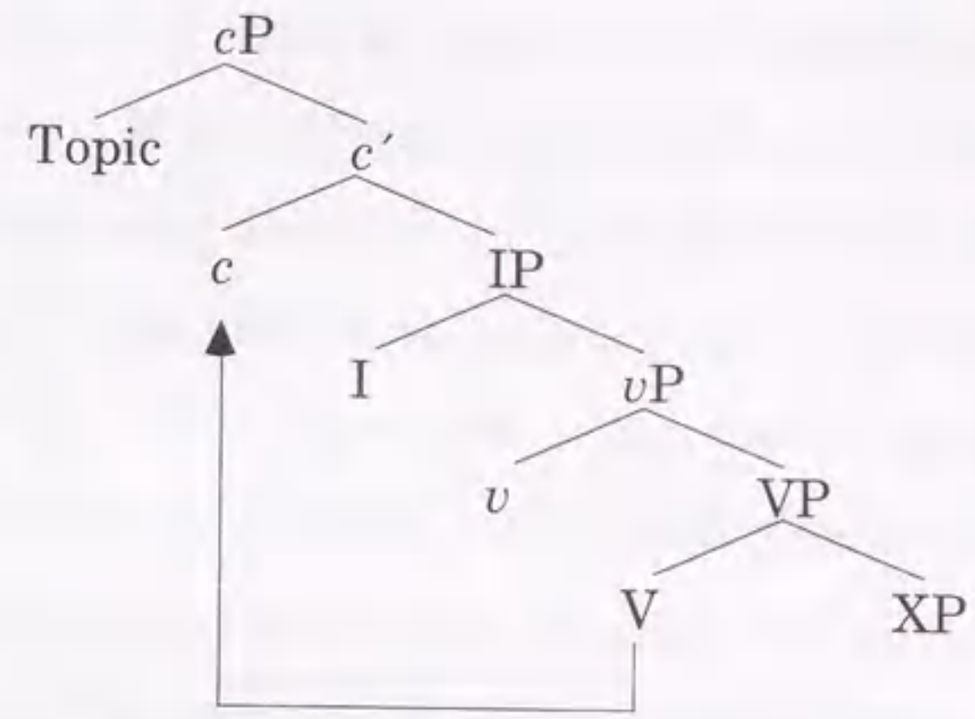
It must be made clear, for the sake of validity of our account, why the option of V2 is available in mainland Scandinavian languages.

Concerning the trigger of V2 movement, two types of proposals have so far been put forth in the literature. On the one hand there are some authors such as Holmberg and Platzack (1995) and Watanabe (1994) who argue that V2 movement is triggered by some attracting feature and takes place independently of XP movement of the topic into [Spec, *c*P], as shown in (67a). On the other hand, Zwart (1993) hypothesizes that V2 movement is indirectly triggered by the topic element in [Spec, *c*P], as shown in (67b). He maintains that V-raising takes place for the purpose of removing features that would otherwise block the topic movement. Under this view, V2 movement is not feature-driven in a strict sense, but it is required to establish a Spec-head relation between the topic and V:

(67) a.



b.



So far, we have taken the former approach, positing [+F] as the attracting feature that induces overt verb movement. This approach has been motivated by the correlation between rich verbal morphology and V2 movement observed in other Germanic languages and valid for explaining the simultaneous demise of V2 and the past plural agreement morpheme in the history of English. However, this is not necessarily to deny the possibility in (67b) in other languages. Even in OE and early ME, it is likely that the interpretation of the sentence initial element as the topic helped children to place [+F] on *c*. Then, I would like to suggest that V2 movement in mainland Scandinavian languages falls under type (67b) and propose the following language-specific condition to explain topicalization in these languages:

(68) A [+Top] feature can enter into a checking relation with a topic only if V[+finite] adjoins to it.

Let [+Top] to be a feature that triggers topicalization. In languages like German and present-day English, this feature does not demand

the support of finite V when it attracts a topic into its checking domain, and some of these languages including German additionally exhibits V2 word order because of the distinct [+F] feature on *c*. On the other hand, V2 in mainland Scandinavian languages is induced by the condition in (68). Movement of this type is analogous to locative and quotative inversions in English mentioned in (32) in that more than one feature participates in single XP movement. Thus, (68) can be regarded as a kind of feature collaboration. The present analysis is indeed consistent with the condition on feature collaboration stated in (31), since [+Top] counts as a (non-logical) operator feature.

It is difficult to make a distinction between the two approaches to V2 movement because either will result in the same output in main clauses. Yet the difference manifests itself in embedded clauses. Compare (69a) and (69b). In German, embedded V2 is totally excluded when the complementizer *daß* appears; on the other hand, in Danish, embedded V2 is compatible with an overt complementizer when it is preceded by certain bridge verbs. The same phenomenon is observed in Swedish and Norwegian as well:

(69) a. *German*

*Sie sagte, $[_{cP}$ daß~~[+F]~~ $[_{CP}$ wir_i $sollten_j$ $[_{IP}$ t_i keine Bücher
she said that we should no books
kaufen t_j]]].
buy
'She said that we should buy no books.'

b. *Danish*

Hun sagde [$_{cP}$ at [$_{CP}$ vi$_i$ skulle$_j$ [$_{IP}$ t_i ikke t_j købe denne
she said that we should not buy this

bog]]].
book

'She said that we should not buy this book.'

(Vikner (1995:84-85))

The ungrammaticality of (69a) naturally follows if the driving force of V2 movement is [+F] in German. Since this feature has already been checked off by *daß*, V-raising cannot be motivated. The complementizer and finite verbs are in a strict complementary distribution so that embedded V2 is only grammatical if there is no *daß*. By contrast, the cooccurrence of the complementizer and embedded V2 in (69b) suggests that V2 movement in Danish does not take place to check off [+F], but to support the topic element in [Spec, Comp]. Note that when there is no topicalization in embedded clauses, finite verbs follow VP adverbs in Danish, as shown in (70):

(70) *Danish*

Peter troede at Helge [$_{vP}$ *gerne* *ville* læse den
Peter believed that Helge readily would read this

her bog].
here book

'Peter believed that Helge would readily read this book.'

(ibid.:143)

Since Danish is a head-initial language (cf. Vikner (1995)), the fact that finite verbs follow VP adverbs indicates that they do not move to

Infl in overt syntax. In terms of the theory presented here, [+F] is located in v in Danish because of poor inflection and a finite verb cannot move any further unless topicalization requires its support in the relevant functional head positions.[26] It can therefore be concluded that V2 movement in mainland Scandinavian languages is not the kind of feature-driven movement discussed in the preceding sections.

Even some languages with rich verbal morphology can employ the strategy in (67b) to derive the V2 word order. Like mainland Scandinavian languages, Icelandic and OE exhibit embedded V2 after overt complementizers, as in (71):

(71) a. *Icelandic*

Jón efast um að *á morgun fari* María snemma
John doubts on that tomorrow will Mary get up
á fætur.
early (=(18a) in Ch.3)
'John doubts that Mary will get up early tomorrow.'

(Vikner (1995:72))

b. *OE*

Gregorius se trahtnere cwæð þæt *forði* *wolde*
Gregory the interpreter said that therefore wanted
drihten getrahtnian þurh hine sylfne þæt bigspel ...
God interpret-Inf through himself the parable
'The interpreter Gregory said that therefore God wanted
to interpret the parable by himself . . . '

(Thorpe *Ælfric's Homilies* ii.88.13, van Kemenade (1997:333))

Since [+F] in *c* is checked off by the complementizers, V must have raised to support the topic elements in [Spec, CP]. Thus, it follows that these languages can resort to the condition in (68) to derive the V2 pattern in embedded clauses.[27]

To sum up, the V2 languages we have reviewed are classified into the following three groups according to the strategies they employ to obtain their V2 patterns:

(72) *V2 Strategies*

a. Type (67a): German, Dutch[28]

b. Type (67b): Swedish, Norwegian, Danish

c. Hybrid: Icelandic, OE

In Icelandic and OE, V2 in main clauses are triggered by [+F] in *c*, whereas in embedded clauses with overt complementizers, V raises to Comp (the lower C) to support a topic element. Thus, they are classified into "hybrid" V2 languages in that both of (67a) and (67b) are available as a means to derive the V2 word order.

4.8. Summary

In this chapter, we have seen how diachronic feature shift occurs in the process of language acquisition by examining the loss of verb movement and the rise of *do*-support in the history of English. Viewing the dummy *do* as an X^0 expletive, I postulated the apparently irregular distribution of [+F]. Then I introduced the notion of Attraction Domain and incorporated it into the definition of Attract F. As long as English verbal morphology retained rich

inflections, the Attraction Domain covered *c*P so that [+F] could be located in *c*. As verbal morphology gradually declined, however, the Attraction Domain reduced to IP or to *v*P, and children exposed to PLD with poor morphology shifted the position of [+F], except for cases where certain lexical elements intervened between neighboring head positions. When [+F] remained outside of the Attraction Domain and no other auxiliaries were employed, *do*-support became obligatory as the last resort to check off [+F]. Thus, the language-specific nature of *do*-support in present-day English can be explicitly understood if we take into consideration the mechanism of parameter resetting in the diachronic dimension.

The analysis made in this chapter has some implication for the theory of language change. In analyzing diachronic feature shift, I have invoked no transcendental principles but maintained that only the decline of verbal morphology is responsible for parameter resetting. This position is contrary to those opinions such as Roberts (1998a), who argues that V-to-I movement is bound to be lost on the grounds that UG prefers those parameter settings which generate relatively simple representations. Under this view, it is not clear at all why V-to-I operations ever developed in grammars (see Chapter 1). On our account, the parameter setting in present-day English is not so simple because (i) [+F] is located in different positions in different types of sentences and (ii) even in a particular type of sentences, the operator sub-feature *[+F]* and the strong sub-feature **[+F]** may be separated. Given our characterization of [+F] as a

feature that takes scope over Infl at LF, *c* is regarded as the default position of [+F], and diachronic feature shift takes place only when it is forced by the decline of verbal morphology. Thus, the locality property of language change observed in the preceding chapter is explained by the theory of UG, together with the general tendency of language acquisition to maximally preserve the surface word order of PLD. This analysis is in accordance with the spirit of the anti-teleological view of language change assumed throughout this thesis, which stresses that there is no transcendental principles that govern the process of parameter resetting.

notes

* This chapter develops in much greater detail the materials presented at the 16th National Conference of The English Linguistic Society of Japan held at Tohoku University on November 7-8, 1998, which was published as Nawata (1999a). I am grateful to Hirozo Nakano, Masachiyo Amano, Masayuki Ohkado, Mitsuru Maeda, Kazuhisa Ishikawa, Hiroshi Terada and Madoka Murakami for their comments and suggestions on earlier versions of this chapter and on the presentation.

[1] For the exposition of previous studies of the auxiliary *do* made under pre-minimalist frameworks and their problems, I refer the reader to Amano (1994) and will not go into details here unless they have some implication for our analysis.

[2] If one adopts Distributed Morphology advanced by Halle and Marantz (1993), it might be argued that phonological features of auxiliaries are not included in the narrow syntax but assigned to them after Spell-Out on their way to the PF output (late insertion). Even if this is true, however, an arbitrary stipulation must still be made that the option of late insertion is not available for

main verbs, which are always pronounced at the original position.

[3] Here, the distribution of [+F] is classified according to sentence types, but I do not wish to imply that sentence types such as affirmative declaratives and negative sentences are primitive notions of UG. They are employed only as mnemonics to describe the positions of [+F]. It is also important to bear in mind that (5) holds only for present-day English. In French, for example, *wh*-interrogatives without subject-verb inversion are allowed.

[4] Rizzi (1991) proposes that Infl in matrix interrogatives is specified as [+wh] and this [+wh] Infl moves to Comp to satisfy his *wh*-criterion. This analysis cannot be maintained under the current framework for the following reasons. First, we are assuming that features which determine clause types including [+wh] are uniformly located in *c*, as discussed in Chapter 2. Secondly, following the spirit of the minimalist program, we interpret the *wh*-criterion as an output condition (cf. note 6 in Chapter 3); thus, it cannot be a trigger for movement by itself and there must be an independent motivation for subject-auxiliary inversion.

[5] I will suggest below that ellipsis is a syntactic operation that applies prior to Spell-Out. Under this view, it will be possible that in (8c, d) VP in the second clauses is deleted before the antecedent *have* in the first clauses raises to Infl. Therefore, the ungrammaticality of these sentences cannot be attributed to the presence of the negative marker *not* in the antecedent phrases.

[6] The progressive auxiliary *was* generated under F^0 in (12) is subcategorized for taking a present participle as its complement. I would assume that after *-ing* adjoins to the stem *make*, its formal features raise to F^0 (or to Infl) in covert syntax to satisfy the selectional property of *was*. This movement may be overt in languages that exhibit V-to-I movement of finite verbs (see (57) in Chapter 3).

[7] Grammaticality of ellipses of this type is not consistent. For example, Lasnik (1995) judges (i) as ungrammatical:

(i) *John [slept], and Mary was ~~sleeping~~ too. (Lasnik (1995:265))

He assumes a structure in which the affix *-ing* takes VP headed by the stem *sleep* as its complement, so that the ungrammaticality of this example can be accounted for by his *stranded affix filter*, which requires that affixes merge with V at PF. However, this analysis fails to account for the fine examples in the text. Then, let us suppose, following Potsdam (1997), that the degraded acceptability of (i) is due to processing difficulty. If semantic information sufficient to recover the elided VP is given, acceptability will improve.

[8] If SOV languages are in fact head-initial, as Kayne (1994) assumes, this word order is derived through overt object shift. See Zwart (1993) and Roberts (1997) for head-initial analyses of Dutch and OE, respectively. I leave the question open here whether the criterion in (14) holds for all SOV languages.

[9] Roberts (1993) proposes a different formulation from Vikner's, which states that V-to-I movement is triggered by the existence of overt distinct number morphology. According to this criterion, it is predicted that verb raising should have been lost as early as in 1500 since distinct number morphology had already eroded in this period. Thus, there would be a gap of about half a century between the loss of inflection and the loss of V-to-I.

[10] In Romance languages, simple past forms illustrated in (18) are mainly employed in written styles. There are two other ways to express past events or state of affairs in conversational styles. Composite past tense is the most usual way to describe past events, and it is represented in the form of *have/be* cognates with present tense + past participles. Imperfect past tense, on the other hand, corresponds to the past progressive in English. Though a common thematic vowel often appears in this form, it may well bear an aspectual meaning. It thus can be said that distinctive past tense morphology does not appear in colloquial styles as well.

[11] The definition of Attraction Domain presented here might seem to be somewhat stipulative, since the domains of each class of affixes are absolutely, not relationally, defined (just as bounding nodes are simply defined as IP and NP for English in the early P & P model). These domains might be reformulated in

terms of Chomsky's (1998) *phase*, a notion which incorporates both *v*P and CP (= our *c*P). If we employ this notion, (26b) could be restated in such a way that Degree I affixes are visible only within the first phase and Degree III affixes are visible until the second phase is closed in the derivation. This is, however, only a speculative idea at present.

[12] On the basis of his split-Infl hypothesis, Pollock (1989) originally proposes that the finite Tense node that is either [+past] or [−past] is an operator. However, since we abstract [+F] away from specific values of Tense (= Infl) and assume that only the former is responsible for finiteness of sentences, the Tense node does not count as an operator under the present framework.

[13] I leave open here the precise position of the pronominal subject in inverted interrogatives. It may be attached to the right of the finite V or be in [Spec, IP], but the choice of these options will not affect our discussion. See van Kemenade (1987) and Hulk and van Kemenade (1997) for further details of this matter.

[14] Of course, when a *wh*-phrase itself is the subject, it precedes the finite verb. This subject/non-subject asymmetry concerning word order has much to do with the asymmetry in the distribution of auxiliaries in interrogative sentences. For specific details, see the following discussion.

[15] Thus, our hypothesis is not consistent with the subset principle for parameter setting proposed by Berwick (1985). For a critical reconsideration of this principle, see Otsu (1989).

[16] Boser et al. argue that when non-finite V appears in the sentence finial position, a null auxiliary is inserted into the Comp position. This is analogous to the dummy *do*-insertion in English.

[17] Shima (1999) also independently develops an idea that the feature of Comp should be divided into a strong *wh*-feature and an operator feature. Thanks are due to Hiroshi Terada for providing me with information on this paper.

[18] Aspectual auxiliaries *have/be* are also dominated by F^0, but unlike modals,

they head their own maximal projection FP and take *v*P (or another FP) as their complement. We will turn to the derivation of sentences involving *have/be* in the next section.

[19] The example in (49a) from the 15th century is a potential problem, since overt V-to-I movement was still prevalent in this period. According to Kroch (1989), however, the adverb-V word order in affirmative declaratives is sporadically observed in the 15th century. Thus, (49a) can also be regarded as a forerunning instance of the Modern English pattern.

[20] In what follows, I will use the simple notation [+F] for the strong sub-feature **[+F]**, unless the distinction is relevant to our discussion.

[21] This assumption is independently needed to deal with long-distance head movement in OE and early ME. When finite V moves to *c* in a successive cyclic way in accordance with the shortest move requirement, it first raises to *v* and the complex head then raises to Infl. Finally, $[_I$ Infl $[_v$ *v* [V[+finite]]]] adjoins to *c*. For [+F] on *c* to be properly checked off, the [+finite] feature most deeply embedded in the complex head must be in the checking domain of *c*.

[22] The condition in (56) might be problematic because movement of bare infinitives is not feature-driven and it is not clear how to incorporate this condition into the system of Attract F. As for LF replacement of *there*, Chomsky (1995:364) assumes that the categorial D-feature of *there* is uninterpretable and must be deleted by some appropriate feature. He then argues that the formal features of an associate, which adjoin to Infl to check off tense and ϕ-features of Infl, carry along the categorial [N] feature as a free rider. The uninterpretable D-feature of *there* attracts [N] to its checking domain and forms a configuration $[_D$ N Exp]. In this configuration, D-feature is properly deleted. Applying this line of reasoning to the present case, we can maintain that the categorial V-feature of *do* is uninterpretable and must be deleted by the interpretable [V] feature of a bare infinitive. Just as *there* requires [N] of an indefinite associate (the definiteness effect), *do* requires [V] of a bare infinitive with semantic contents.

[23] As Quirk et al. (1985:827) note, the perfective form is rare in imperative sentences. This is most plausibly because imperative sentences are primarily used to refer to irrealis events.

[24] Kazuhisa Ishikawa (personal communication) points out that *-ing* nominals such as *bringing-up*, which are derived from verb-particle combinations, may raise problems for the present analysis. If the complex verb [$_V$ bring up] is listed in the lexicon and the affix -ing is attached to this complex, the resulting structure will be [$_N$ [$_V$ bring up] -ing]. In this structure, *bring* and *-ing* is not adjacent, so that *bringing-up* must be derived by means of a morphological rule that directly attaches *-ing* to *bring*, ignoring the word-internal structure. Let us tentatively assume that this kind of morphological reanalysis is limited to lexical operations and cannot be applied to verbal complexes formed by syntactic operations; otherwise, the affixal property of -IMP would be satisfied in the configuration [$_I$ [$_{Neg}$ have/be not] -IMP], contrary to fact.

[25] There are many other complexities about *do*-support in imperatives. For example, when an overt subject appears in negative imperatives, *don't* precedes the subject, as in (i):

(i) a. Don't you sit down!
b. *You don't sit down!

In this case, we must suppose that *don't* raises to Comp. However, Potsdam (1996) argues that Infl-to-Comp movement of *don't* is optional by giving the following examples:

(ii) a. Those with luggage don't leave it unattended!
b. Girls go into the hall, BOYS don't move! (Potsdam (1996:387))

What factors are involved in this optional movement is a question which I want to keep beyond the scope of the present discussion. See Potsdam (1996) for a fuller study of relevant issues of imperative *do*.

[26] Mainland Scandinavian languages underwent the loss of V-to-I movement in embedded clauses between 1500 and 1800, along with the loss of verbal

morphology (cf. Vikner (1995, 1997) and references cited therein). Thus, these languages are no exceptions to the generalization that only languages with relatively rich verbal morphology exhibit V-to-I movement.

[27] In Icelandic, V2 can occur freely in embedded clauses whereas in OE embedded V2 is allowed only after certain bridge verbs. This difference can be reduced to the morphological properties of complementizers in these languages. See (19) in Chapter 3 for the structure of general embedded V2 and (37) in Chapter 2 for the structure of restricted (exceptional) embedded V2.

[28] Vikner (1995:66) notes that embedded V2 is not possible in Dutch for some unknown reason.

Chapter 5

Diachronic Feature Shift in Infinitival Clauses

5.1. Introduction

In the preceding chapters, I have demonstrated that the development of the complementizer system and the loss of verb movement in the history of English are properly analyzed as effects of the diachronic feature shift of [+F] among functional categories. However, since these changes all occurred in finite clauses and we have primarily concentrated on the shift of [+F], it is now necessary to show that the present theory can be justified in the case of non-finite clauses as well. The aim of this chapter is to extend the idea of diachronic feature shift to infinitival clauses and demonstrate that this mechanism has a wide range of explanatory power.

More specifically, this chapter presents an analysis of the historical transitions of three types of English infinitival constructions: the accusative with infinitive (ACI) construction, the *for* NP *to* VP construction, and *for to* infinitives. It will be argued that inflected infinitives in OE lacked independent Infl node and that

their structure changed during ME along with the decline of the infinitival affix *-enne*. First, in the mid-14th century, *to* was reanalyzed as a tense (Infl) element and consequently infinitives became more articulated than in OE. I will claim that this was the immediate cause of the spread of the ACI to *believe*-type mental perception verbs and the rise of the *for* NP *to* VP construction. Then later in the mid-16th century, the infinitival affix was completely lost, and this led to the demise of *for to* infinitives. Contrary to finite clauses, infinitival clauses have no [+F] feature because their temporal interpretations are always dependent on matrix finite clauses; instead, the infinitival affix *-enne* in OE had an interpretable [−finite] feature, which shifted from F^0 dominating the affix to the newly introduced Infl node. Thus, the diachronic feature shift in infinitival clauses is an upward shift, unlike the downward shift of [+F] in finite clauses.

The organization of this chapter is as follows. Section 5.2 lays the groundwork for a historical analysis by establishing the structures of infinitival complementizers in present-day English. Section 5.3 reviews basic facts about the history of the infinitival constructions and points out some problems with previous analyses. Section 5.4 examines the structural and functional changes of infinitive markers and gives an explanation for the historical transitions of *for to* infinitives and the *for* NP *to* VP construction. Section 5.5 considers the correspondence between functional categories and semantic features in the ACI construction.

5.2. The Structures of Infinitival Complementizers

In Chapter 2, I developed the split-CP hypothesis and proposed that complementizers employed in finite clauses should be decomposed into *c* and C. The possible combinations and their morphological realizations are repeated here for the sake of reference:

(1) *Finite Complementizers*

[−wh] *c* +C (subordinator) : that

[+wh] *c* +C (subordinator) : whether

As long as this approach is on the right track, the same analysis will naturally hold for infinitival complementizers. *Whether* is employed to introduce infinitival clauses as well as finite clauses. In addition, I propose that the complementizer *for*, like *that* in finite clauses, should be decomposed into [−wh] *c* and the subordinator C:

(2) *Infinitival Complementizers*

[−wh] *c* +C (subordinator) : for

[+wh] *c* +C (subordinator) : whether

As discussed in Chapter 2, each head of the split-CP structure has nothing but syntactic or semantic features (clause type specifications of *c* and the subordinator property of C), and phonological features are assigned at the PF side to the amalgamate formed either by overt raising of C into *c* or by morphological merger operation under structural adjacency between these heads.

A significant lexical property of the infinitival complementizer

for is its Case-checking ability. Under the split-VP hypothesis, the ability of a verb to check Case is attributed to the upper v, which is included in transitive and unergative verbs but not in unaccusative verbs (Chomsky (1995:315-316)). If a parallel argument can be made with respect to the split-CP structure, the upper c will be responsible for the Case-checking ability of *for*. Then, let us make the following assumption:

(3) In infinitival clauses, [−wh] c has a Case-assigning feature.

Note, however, that since the Case feature of the subject is checked off by Infl in finite clauses, c does not have a Case feature when it is a constituent of *that* or *whether*. In that case, c serves as a pure clause-type indicator without any Case to assign.[1]

The proposed system has an advantage of being able to distinguish between two kinds of null Comps in infinitival clauses. Let us consider the following well-known contrast:

(4) a. John tried [ϕ [PRO to leave]].

b. *John tried [ϕ [him to leave]].

c. John wanted [ϕ [PRO to leave]].

d. John wanted [ϕ [him to leave]].

When an infinitival clause lacking *for* occurs as the complement of *try*-class verbs, the subject must be PRO, and no lexical subject can appear as shown in (4a, b). This suggests that the null Comp ϕ has no Case-checking ability.[2] On the other hand, *want*-class verbs allow a lexical subject as well as PRO to appear in their infinitival

complements, as illustrated in (4c, d). Lasnik and Saito (1991) convincingly argue that *him* in (4c), unlike the subject of infinitives in the *believe*-type ECM construction, does not undergo A-movement into the matrix clause to get its Case checked off by the verb. In (5a), the embedded subject cannot corefer with the referring expression within the matrix adverbial clause, whereas the same coreference does not yield total ungrammaticality in (5b):

(5) a. ?*Joan believes [him$_i$ to be a genius] even more fervently than Bob's$_i$ mother does.

b. ?Joan wants [him$_i$ to be successful] even more fervently than Bob's$_i$ mother does. (Lasnik and Saito (1991:336))

This indicates that *him* in (5a) c-commands *Bob* at least covertly from the matrix Case-checking position and hence causes a violation of the Condition C of the Binding Theory, while the subject in (5b) stays in the embedded clause throughout the derivation and the Condition C is properly observed. Likewise, the subject of ECM infinitives can license anaphors and negative polarity items in the matrix clause, whereas the subject of *want*-class infinitives cannot:

(6) a. ?I believed [those men$_i$ to be unreliable] because of each other's$_i$ statements.

b.??*I wanted [those men$_i$ to be fired] because of each other's$_i$ statements. (ibid.:337)

(7) a. ??I believed [none of the applicants to be qualified] after reading any of the reports.

b.??*I wanted [none of the applicants to be hired] after reading any of the reports. (ibid.)

These facts also suggest that only the embedded subjects of the ECM infinitives undergo A-movement into the matrix clause for Case-checking and c-command the anaphors and the negative polarity items from the landing site. It follows that the Case feature of the subject of *want*-class infinitives is not checked off against the matrix verb. It is also clear that there is no inherent Case-marker within infinitives. Therefore, the only element that can Case-check the subject of *want*-class infinitives is the null Comp ϕ. The question at issue, then, is why the null Comp ϕ in *try*-class infinitives cannot check the Case of the embedded subject unlike its counterpart in *want*-class infinitives.

In his analysis of the *for* NP *to* VP construction, Watanabe (1993b) proposes that the *for-to* complex is base-generated under the Infl node and the subject is Case-checked in [Spec, IP] under Spec-head agreement with the *for-to* complex, and that *for* subsequently raises to Comp to derive the surface word order. Essentially following this analysis, Bošković (1997) argues that the null Comp ϕ in *want*-class infinitives behaves like *for* in all relevant respects. Thus, the derivation would be illustrated as follows:

(8) John wanted [[him for/ϕ-to leave]]
~~[Case]~~

As for the driving force of the overt movement of *for* and its null counterpart, Bošković assumes that they raise to Comp to satisfy the

selectional property of the matrix verb.

However, the Watanabe-Bošković analysis of *want*-class infinitives faces some conceptual and empirical problems. Conceptually, there seems to be no necessity of generating *for* under the Infl node. Bresnan (1972) argues that the complementizer *for* is semantically active, on the basis of the observation of Kiparsky and Kiparsky (1970) that *for* occurs only with emotive predicates and expresses meanings like subjective reason, purpose, or goal. Since these meanings are closely related to clause type specifications rather than tense or agreement, it seems more reasonable to think that *for* is directly generated under Comp. A more serious empirical problem concerns the distribution of the null Comp ϕ. Both *for* and its null counterpart can appear in the complement of *want*, but only *for* is allowed when infinitival clauses occur in the subject position and a non-subcategorized adjunct position, as illustrated in (9):

(9) a. [For [him to submit]] would be impossible.

b. *[ϕ [him to submit]] would be impossible.

c. We left it on the porch [for [Mary to find]].

d. *We left it on the porch [ϕ [Mary to find]].

Moreover, if *for* is base-generated under Infl and raises to Comp to satisfy the selectional property of the matrix verb, it is predicted that *for* may stay in situ when the infinitival clause does not occur in a subcategorized position. This prediction, however, is not borne out:

(10) a. *[Him for-to submit] would be impossible.

b. *We left it on the porch [Mary for-to find].

The Case feature of the embedded subjects would be properly checked off under Spec-head agreement with the *for-to* complex. The ungrammaticality of these examples strongly suggests that it is implausible to say that *for* and *to* form a complex in the underlying structure at least in present-day English.[3] Thus, we need a conceptually desirable theory that does not generate *for* and ϕ under Infl which also explains the distribution of the null Comp ϕ in (9).

Our split-CP hypothesis achieves this result. I assumed above that *for* is a morphological realization of the amalgamate [−wh] *c* +C and its Case-checking ability is attributed to the upper *c*. In this system, Case-checking of the subject of infinitives is not carried out under Spec-head agreement within IP, but through feature movement in covert syntax:

(11) $[_{cP}\ c(-\text{wh})\ [_{CP}\ C\ [_{IP}\ \text{Subj}\ \text{to}\ \text{VP}]]]$
~~[Case]~~
FF(Subj)

After Spell-Out, [−wh] *c* attracts the formal features of the subject including Case and gets its Case-assigning feature checked off in the narrow syntax, while *c*+C is spelled out as *for* at the PF side. If an infinitival clause lacks either *c* or C, *for* is not morphologically realized, so that the null Comp ϕ will be obtained:

(12) [−wh] *c* without C
C without [−wh] *c* $\rangle$ ϕ

In the present framework, therefore, infinitival clauses involving *for* and ϕ have different structures and there are two kinds of null infinitival Comps: one is simple *c* without C and the other is simple C

without *c*, as illustrated in (12). Of these two kinds of null Comps, only the former has a Case-checking ability.

Assuming this much, let us return to the contrast between *try*-class verbs and *want*-class verbs in (4). It was postulated in the standard Government-and-Binding Theory that these verbs all select CP in the traditional sense and only infinitival complements of *want*-class verbs contain the null Comp ϕ with a potential Case-assigning ability, while the head C of *try*-class infinitives is literally empty. This assumption, however, cannot be maintained in the current minimalist model, because an empty head without any effects on both interface levels, LF and PF, does not enter the numeration in view of economy considerations (Chomsky (1995:294)). Under the split-CP hypothesis, the contrast between the two classes of infinitival complements can be reduced to the selectional properties of the matrix predicates, given the two structurally differentiated null Comps in (12).[4] Let us assume, then, that *try*-class verbs select only CP whereas *want*-class verbs select *c*P in addition to CP:

(13) a. try: _____CP

John tried [$_{CP}$ C [$_{IP}$ PRO to leave]].

b. want: _____*c*P, CP [5]

John wanted [$_{CP}$ C [$_{IP}$ PRO to leave]].

John wanted [$_{cP}$ *c*(−wh) [$_{IP}$ him to leave]].

covert c*-incorporation*

Since these null Comps have syntactic or semantic features of their own, they have interpretive effects at LF, so that they are not

excluded from the numeration. When a *try*-class verb selects CP, the embedded subject must be realized as PRO because the head C has no Case feature. *Want*-class verbs, by contrast, have another option of selecting *c*P. Though *for* is not morphologically realized when *c*P lacks CP and immediately dominates IP, the embedded lexical subject is nevertheless licensed owing to the Case-checking ability of *c*. After Case-checking, *c* in (13b) covertly incorporates into the matrix verb to get the subordinate interpretation just as in the *that*-less clause discussed in Chapter 2 (see (24) in p.39).[6]

The distribution of the null Comp ϕ in (9) is also naturally accounted for. Elaborate structures of legitimate and illegitimate infinitival clauses in non-subcategorized positions will be represented as follows:

(14) a. $[_{cP}$ For $[_{CP}\, t_C\, [_{IP}$ him to submit]]] would be impossible.

b. *$[_{cP}\, c(-\text{wh})\, [_{IP}$ him to submit]] would be impossible.

c. *$[_{CP}\, C\, [_{IP}$ him to submit]] would be impossible.

d. $[_{CP}\, C\, [_{IP}$ PRO to submit]] would be impossible.

(15) a. We left it on the porch $[_{cP}$ for $[_{CP}\, t_C\, [_{IP}$ Mary to find]]].

b. *We left it on the porch $[_{cP}\, c(-\text{wh})\, [_{IP}$ Mary to find]].

c. *We left it on the porch $[_{CP}\, C\, [_{IP}$ Mary to find]].

d. We left it on the porch $[_{CP}\, C\, [_{IP}$ PRO to be found]].

When an infinitival clause contains a lexical subject, *c* must appear to check off the Case feature of the subject. If Comp has a full-fledged *c*P-CP structure as in (14a) and (15a), *c*P is properly interpreted as a subordinate clause by the presence of the

subordinator C at LF and the amalgamate of the two heads is morphologically realized as *for* at PF. Since *for* is a genuine complementizer, the Subj-*for*-*to*-V word order in (10) does not simply occur. On the other hand, if *c*P alone occurs as in (14b) and (15b), *c* must covertly incorporate into the matrix verb to receive the subordinate interpretation, but *c*-raising into V is impossible from these positions. Furthermore, if CP alone occurs as in (14c) and (15c), the Case feature of a lexical subjects will not properly be checked off. The distribution of the null Comp ϕ with a Case-checking ability is thus limited to the complement position of V. By contrast, when the subject of an infinitive is PRO as in (14d) and (15d), *c* is not required and only CP appears. Since C itself has a semantic feature responsible for the subordinate interpretation, it need not incorporate into the matrix verb, so that simple CP infinitives without lexical subjects can appear in any position. In this case, they are regarded as [−wh] through default interpretive procedure.

From the discussion above, it can be concluded that there is good reason to assume the split-CP structure for the infinitival complementizer *for* as well as the finite complementizers.

5.3. Basic Facts and Previous Analyses

5.3.1. Chronological Digest

Now that the theoretical background has been established, let us review some basic facts about the historical development of

infinitival constructions. In OE, only causative verbs and perception verbs could take the accusative with infinitive (ACI) complement. As illustrated in (16), they were followed by bare infinitives:

(16) a. se . . . doeð hie cwacian.
that makes them quake-Inf
'that . . . makes them quake'
(*The Vespasion Psalter* 103.32,Visser (1963-1973:2257))

b. Ic seah turf tredan.
I saw turf tread-Inf
'I saw the turf trod.' (*Riddles* 13.1, ibid.:2254)

The ACI construction with *believe*-type mental perception verbs was not prevailing in OE. Although a few examples are actually adduced in the literature, they are all considered to be strongly influenced by Latin originals as argued by Fischer (1989, 1992). True ACI complements first appeared with *believe*-type verbs in late ME:

(17) I beleeue euerlasting liif to be
I believe everlasting life to be (*Donet* 104.7, ibid.:2309)

Before this construction came into use, *believe*-type verbs could take *that*-clauses, small clauses, and control infinitives as their complements. They are exemplified in (18):

(18) a. That-*Clause*
swa ðæt he secge oððe gelyfe ðæt þry Godas syndon
so that he say or believe that three Gods are
'so that he might say or believe that there are three Gods'
(Bosworth *John* 14.11, *An Anglo-Saxon Dictionary*)

b. *Small Clause*

þa þing Op$_i$ þe ge . . . t_i betstan gelyfað
the things that you best believe

'the things that you believe to be best'

(*Bede* 1.25, Visser (1963-1973:558))

c. *Control Infinitive*

Ic gelyfe to arisenne on domes dæge.
I believe to rise on doom's day

'I believe that I will rise on Doomsday.'

(Thorpe *Ancient Laws* ii.262.15, ibid.:1323)

(18c) is now obsolete. In what follows, the term "B-type verbs" will be employed to refer to these mental perception verbs.

Another new pattern that developed in late ME was the *for* NP *to* VP construction. The earliest example of this construction in Visser (1963-1973) is dated 1380, which coincides with the spread of the ACI construction to B-type verbs:

(19) make ȝe redi a hors for poul to ride on.
make you(sg.) ready a horse for Paul to ride on

(1380 Wyclif, *Acts* 23.24, Visser (1963-1973:988))

On the other hand, *for to* infinitives without lexical subjects like (20a) came into use as early as in the late 11th century:

(20) a. Godes gerichtten *for to setten* . . . ,
Gods ruled for to set

(1066 *Codex Diplomaticus* IV 306.3, Mustanoja (1960:514))

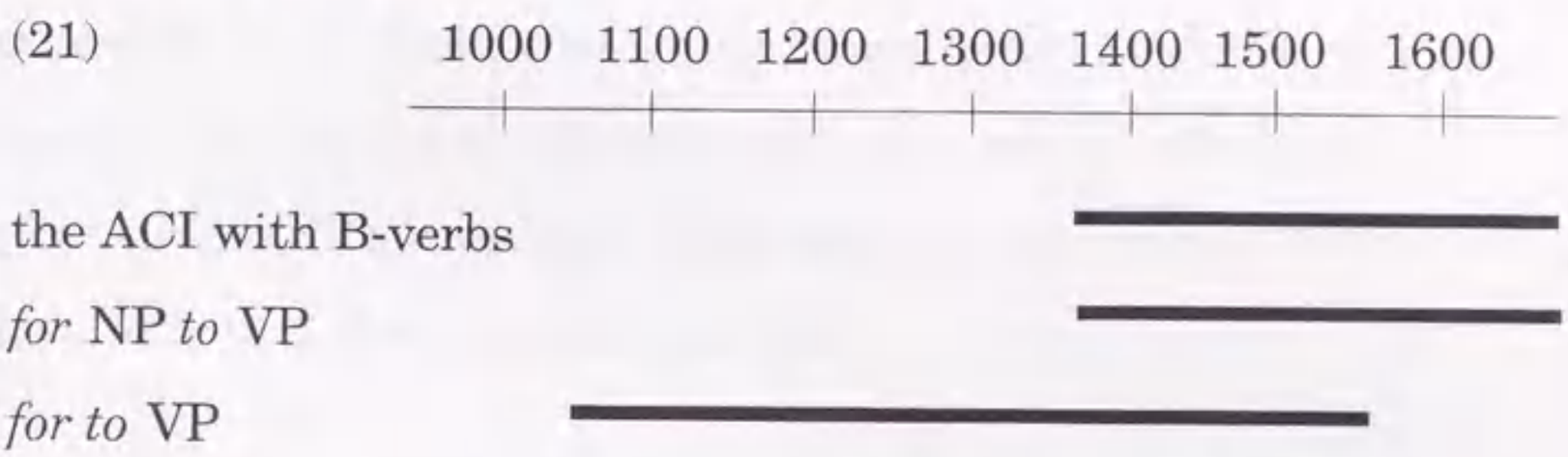
b. So we consulted *for me to go* first to Sir H. Bennet.

(17c *Pepys's Diary* Oct.30, Visser (1963-1973:2246))

However, the frequency of this construction gradually decreased and finally it became obsolete in the mid-16th century, except for certain fixed expressions (cf. Mustanoja (1960), Lightfoot (1979)). Since then, lexical subjects have been obligatory with the complementizer *for*, as in (20b).

Thus, the chronological transitions of the ACI construction with B-type verbs, the *for* NP *to* VP construction, and *for to* infinitives can be summarized as follows:

(21)



A comment is in order here with respect to the status of *for* in the early *for* NP *to* VP construction. Lightfoot (1979) classifies data in Visser (1963-1973) and observes that examples of the following types appeared in late ME:

(22) a. It is necessary for a man to go.

b. I'm afraid for them to see it.

c. This left room for the controversy to go on.

Fischer (1988) and Roberts (1993) argue that *for* NP in early examples of these types was a benefactive complement and they assign them the structure [*for* NP] [PRO *to* VP]. If this is correct, then the "true" *for* NP *to* VP construction that is analyzed as [*for* [NP *to* VP]] should be taken to be an innovation in a later period, perhaps

in Modern English. However, an important point to note is that if *for* in (22) was a benefactive preposition and did not constitute part of infinitival clauses, it is predicted that *for* NP *for to* VP should have also been permissible, since the infinitive markers *to* and *for to* were interchangeable in late ME. As for the type of (22a), Visser lists two examples of *for* NP *for to* VP, but there are no such examples listed with respect to the types of (22b) and (22c), nor have I found any examples of them in the diachronic part of the Helsinki Corpus. This indicates that *for* in these constructions was part of infinitival clauses, not an independent benefactive preposition. We may therefore conclude that at least (22b) and (22c) are true *for* NP *to* VP constructions that developed in late ME.

5.3.2. Previous Analyses

The chronological table in (21) tells us that the rise of the *for* NP *to* VP construction does not coincide with the demise of *for to* infinitives, but there is a period in which both of these options were available in the grammar of English. It has often been argued in the literature that the historical changes of these constructions were triggered by a single parameter resetting. For example, Lightfoot (1979) claims that the emergence of the *for* NP *to* VP construction and the demise of *for to* infinitives were both related to the introduction of the S-bar deletion rule. Roberts (1993) argues that they were triggered by the reanalysis of *to* as a member of Tense. However, these analyses cannot explain why there is a timelag of 200

years between the two changes. It would be unreasonable to assume that a single reanalysis of a grammar took such a long time. To resolve this problem, a more articulated analysis of the historical development of English infinitives must be made.

As for the spread of the ACI construction, Fischer (1989, 1992) discusses it in full detail. She argues that the change of the word order from OV to VO was the immediate trigger for the rise of the ACI with B-type verbs. She points out the fact that *let* in early ME could take the construction in (23b) as well as that in (23a):

(23) a. I prai the let me noght mistime Mi schrifte, . . .
I pray you let me not mis-time-Inf my confession
'I pray you let me not mis-time my confession, . . .'
(*Confessio Amantis* 1.220-221, Fischer (1992:31))

b. heo nom Æstrild & Abren; & lette heom ibinden
they took Æstrild and Abren and let them bind-Inf
'they took Æstrild and Abren and had them bound'
(*Brut* 1244, ibid.)

These constructions differ in the status of the NPs that follow *let*: while *me* in (23a) is the subject of *mistime*, *heom* in (23b) is the object of *ibinden*. Their structures are delineated as follows:

(24) a. let [Subj V ···] (subject construction)
b. let [Obj V ···] (object construction)

Fischer refers to the former type as subject construction and the latter as object construction. The point is that the object construction is a direct reflection of the base OV word order in early

ME; thus, as the word order changed from OV to VO, the object construction also changed from (24b) to (25):

(25) let [V Obj …]

This new pattern, however, was not actually very common, because syntactic subjects became increasingly obligatory in ME. To express the logical meaning of the object construction without recourse to the word order in (25) which lacks the syntactic subject, passive infinitives like (26) came to be preferred:

(26) þa lette he his cnihtes; dœies & nihtes
then let he his knights day-Gen and night-Gen
œuere beon iwepned;
ever be-Inf weaponed
'then he caused his knights always to be armed, day and night' (*Brut* 8155-8156, ibid.:43)

In this example, the underlying object of the infinitive appears in the surface subject position. This means that infinitival complements of causatives were unified into the subject construction.

Fischer claims that passive infinitives with causative verbs spread to other classes of verbs by analogy. First, *persuade*-type three-place verbs, which were semantically similar to causatives, came to occur with embedded passive infinitives as in (27a). Then, she argues that B-type verbs also acquired the ability to take passive infinitives as in (27b) on the grounds that they had two-place, causative-like argument structures:

(27) a. Whan the recluse herde his name she had grete joy of
when the recluse heard his name she had great joy of

hym . . . And than she commaunded the gatis to be
him and then she commanded the gates to be

opyn
open[ed?] (*Morte Darthur* 905.10-13, ibid.:58)

b. she dare not aventure her money to be brought vp to
she dare not adventure her money to be brought up to

London for feere of robbyng . . .
London for fear of robbing

(*The Paston Letters* 156.7-10, ibid.:62)

Thus, Fischer observes that early ACI complements with B-type verbs were strongly influenced by passive infinitives after causative verbs like *let*, which was a by-product of the word order change from OV to VO.

It is indeed very likely that the word order change provided groundwork for the spread of the ACI construction, but Fischer's analysis is not without difficulties. The most problematic one is that the change of word order and the spread of the ACI construction to B-type verbs did not take place simultaneously. According to van Kemenade (1987), the shift from OV to VO completed around 1200. On the other hand, B-type verbs began to take the ACI complement productively around 1400 (cf. van Gelderen (1993)). If the change of word order had been the direct trigger for the spread of the ACI construction, the effect should have manifested itself much earlier. It is unlikely that it took 200 years to introduce a new construction by analogy. Furthermore, the process of the spread of the ACI from causatives to other classes of verbs is not clear. Fischer argues that

passive infinitives with causatives spread to other verbs, but as we have seen in (16), OE causative verbs were already able to take the ACI construction, and B-type verbs could take small clauses. These predecessors could have been potential triggers for the spread of the ACI, and this construction could have been permissible with B-type verbs as early as in OE. From these grounds, it can be reasonably said that the effect of the word order change on the spread of the ACI to B-type verbs was only indirect and there is no direct correlation between them. In the subsequent sections, I will provide a more principled account of the changes illustrated in (21).

5.4. The Historical Development of *To* Infinitives

In this section, I will consider the transitions of the *for* NP *to* VP construction and *for to* infinitives with special reference to the structural development of *to* infinitives in the history of English, delaying an analysis of the spread of the ACI to section 5.5. I will first establish the structure of infinitives in OE and then investigate how it has come to take the shape as it is in present-day English.

5.4.1. The Structure of OE Infinitives

It will be useful to investigate the structure of OE infinitives by focusing on the following two points: (i) the categorial status of the infinitive marker *to* and (ii) the position of the [−finite] feature.

As for the first point, Visser's (1963-1973:976) observation that there were no examples of *wh* + *to* infinitives in OE offers the key.[7]

It is noteworthy that the prohibition against cooccurrence of a *wh*-phrase and an infinitive marker is found in present-day Romance languages as well:

(28) a. *French*

*Je lui ai dit ou d'aller.
I him have told where to go
'I have told him where to go.'

b. *Italian*

*Gli ho detto dove di andare.
him have told where to go
(identical with (28a)) (Kayne (1984:104-105))

Kayne (1984) argues that the French and Italian infinitive markers *de* and *di* are complementizers and that the ungrammaticality of the examples above can be attributed to a violation of the doubly-filled Comp filter in the sense of Chomsky and Lasnik (1977). As discussed in Chapter 3, the doubly-filled Comp filter might be reduced to Rizzi's (1991) *wh*-criterion, repeated here for the sake of reference:

(29) *The* Wh-*Criterion*

A. A *wh*-operator must be in a Spec-head configuration with X^0(+wh).

B. An X^0(+wh) must be in a Spec-head configuration with a *wh*-operator. (cf. Rizzi (1991))

The ill-formedness of the examples in (28) suggests that *de* and *di* are [−wh] infinitival complementizers which cause feature mismatch against the *wh*-operators occupying [Spec, *c*P].

If the infinitive marker *to* in OE was a complementizer like

French *de* and Italian *di*, then the structure of *wh* + *to* infinitives will be represented as follows:

(30) *[$_{cP}$ Wh to(−wh) [$_{CP}$ t_C [$_{IP}$ ⋯]]]

Since *to* infinitives could freely occur in non-interrogative environments, it seems reasonable to suppose that *to* was specified as [−wh]. In (30), the *wh*-phrase is not in a Spec-head relation with an X^0 with a [+wh] feature, so that this configuration leads to a violation of clause A of the *wh*-criterion. Thus, the absence of *wh* + *to* infinitives in OE can be accounted for on a par with the ungrammaticality of sentences like **I do not know where for him to go* and serves as evidence that the OE infinitive marker *to* was a complementizer corresponding to *for* in present-day English.

We may note, in passing, that the complementizer *to* in OE retained its original function as a preposition as well. This is clearly indicated by the dative affix -*enne* attached to the verbal stem following *to*. When infinitives expressed directional or purposive meanings, *to* infinitives were obligatorily selected, but bare infinitives were never. Furthermore, inflected infinitives in OE could be conjoined with full PPs:

(31) Ut eode *to his gebede* oððe *to leornianne* mid his
out went to his prayer or to learn with his
geferum.
comrades
'He went out to give his prayer or to study with his comrades.' (*Bede* 162.7, Kageyama (1992:99))

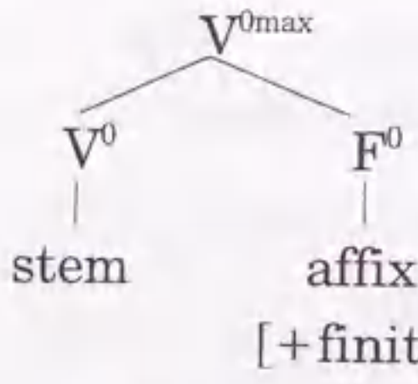
From these facts, it has sometimes been proposed in the literature

that the infinitive marker *to* was a pure preposition in OE (cf. Tanaka (1994, 1997)). This idea is very attractive, but it is also clear that OE *to* infinitives behaved as clauses in light of the fact that they could occur as complements of B-type verbs as in (18c). Then, I regard *to* as a *prepositional complementizer* and label it as *c*+C, just like *for* in present-day English discussed in section 5.2. The parallelism between complementizers and prepositions might be captured in terms of Grimshaw's (1991) notion of "extended projection". Grimshaw defines a functional category as an extended projection of the lexical category that shares the same categorial features. For example, Infl and Comp are the extended projections of V, which is specified as [+V, −N], and D and P are the extended projections of N specified as [−V, +N]. Under this framework, Infl and D are labeled as {F1} and C and P are labeled as {F2} according to the level of projections. Thus, complementizers and prepositions belong to the same level of extended projections. Since OE infinitive verbs with the dative affix *-enne* retained both verbal and nominal properties, the ambiguous status of the infinitive marker *to* between a complementizer and a preposition is not surprising at all. The directional and purposive meanings of *to* can be attributed to the clause type indicator *c* in the same way as *for* in present-day English.

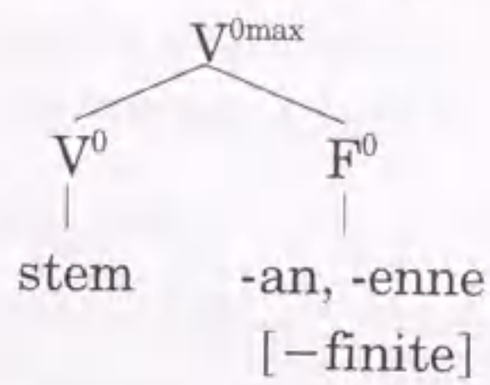
Let us turn to the second issue, the position of the [−finite] feature. In Chapter 4, we argued that finite main verbs have the structure in (32a), in which an affix is dominated by the functional head F^0 that bears the [+finite] feature. It must be noted here that

OE infinitives had independent morphological endings (*-an* for bare infinitives and *-enne* for inflected infinitives) which distinguished them from finite forms. Extending the analysis of finite main verbs to OE infinitives, we obtain the structure in (32b), in which the [−finite] feature is carried by the infinitival affixes:

(32) a. *Finite Verbs*



b. *Infinitives (OE)*



There is, however, a crucial difference between finite verbs and infinitives. It was observed in the preceding chapter that the [+finite] feature is attracted by the corresponding [+F] feature in Infl, whereby a state of affairs expressed by a finite verbs is mapped into some point of the time axis *in relation to the time of utterance* (i.e. independent reference). On the other hand, state of affairs expressed by infinitives are always interpreted *in relation to the matrix tense* (i.e. relative reference), so that there is no need to posit a [+F] feature that attracts the [−finite] feature and forms an operator-variable configuration as in finite clauses. Since the [−finite] feature is interpretable, it need not be checked off, either. Given that there is no [+F] in infinitival clauses and the infinitive marker *to* was a complementizer, it follows that OE infinitives lacked the Infl node altogether, because a category without any contents cannot be introduced into the computational system in the first place

under the current minimalist model.

To demonstrate that the [−finite] feature was carried by infinitival affixes, I would shed light on interpretive ambiguity of causative constructions. In OE, both direct and indirect causatives were expressed by bare infinitival complements. Consider the following examples:

(33) a. & he læt rinan ofer þa rihtwisan . . .
and he causes rain-Inf over the righteous
'and he makes it rain on the just . . .'

(Skeat *Matthew* 5.45, Denison (1993:190))

b. Swa swa ðu dydest minne broðor his god forlætan . . .
as as you made my brother his god forsake-Inf
'Just as you made my brother forsake his god . . .'

(Thorpe *Ælfric's Homilies* i.31.468.20, ibid.:172)

In (33a), *he* exerts his force directly on the nature, so that there is only one event involved in this sentence. On the other hand, the causation in (33b) is indirect in that there is an intermediate Agent *minne broðor* in the causal relation and that the causing event and the forsaking event are not necessarily simultaneous. Thus, it can be said that two distinct events are involved in (33b).

In order to distinguish between direct and indirect causatives, I essentially adopt Higginbotham's (1985) mechanism for licensing event feature specifications of verbs and assume as follows:

(34) The event specification of a verb is anchored to a specific reference time if its event feature is discharged to a tense element.

In the present framework, qualified candidates for the tense element to which an event feature is discharged are [+F] and the [−finite] feature: in finite clauses, an event feature of a verb is discharged to [+F] when the verb is attracted by [+F], and the event is interpreted in relation to the time of utterance; in infinitival clauses, on the other hand, an event feature has another option of being discharged to the [−finite] feature of the infinitive and indirectly interpreted in relation to the matrix tense.

Given this assumption, the relations between event feature specifications and tense elements in (33a, b) can be represented as in (35a, b), respectively:

(35) a. [$_{IP}$ he Infl læt [$_{VP}$ rinan ofer þa rihtwisan]]

<E> <E>

(direct causation)

b. [$_{IP}$ ðu Infl dydest [$_{VP}$ minne broðor his god forlæt-an]]

<E> <E>

(indirect causation)

<E> stands for an event feature of a verb. In (35a) , the <E> features of *læt* and *rinan* are both discharged to the matrix Infl and are combined into a single complex event, resulting in the interpretation of direct causation. By contrast, in indirect causatives where two independent events are involved, the <E> feature of the embedded infinitive must be discharged to a tense element distinct from the matrix Infl. In (35b), the <E> features of *forlætan* is licensed by the infinitival affix, so that the interpretation of indirect causation

results. Thus, the availability of indirect causative reading in bare infinitival complements shows that the infinitival affix -*an* had the ability to bear the [−finite] feature to which an event feature can be discharged. Although the discussion here is based on bare infinitives, it seems quite natural to extend this analysis to inflected infinitives as well, because there is no conceptual reason to suppose that the word-internal structure of V-*enne* is different from V-*an*.

To sum up the discussion so far, it can be concluded that in OE, (i) the infinitive marker *to* was a (prepositional) complementizer, and (ii) the [−finite] feature was carried by the infinitival affix -*enne*. It follows from these two points that OE infinitives had no independent Infl node: since *to* was a complementizer, it could not have been a member of Infl; furthermore, since the [−finite] feature need not be licensed by [+F], there is no need to posit another tense element within infinitives. The structure of OE infinitives can therefore be delineated as follows (the upper *v*P projection is omitted in the following representations unless it has any relevance for our discussion):[8]

(36) *OE*

$[_{cP}\ c(-\text{wh})\ [_{CP}\ C\ [_{VP}\ \text{V-enne}\ ]]]$

to

Like *for* in present-day English, *to* in OE was a complex category made up of *c* and C. Note, however, that *to* was obligatory in inflected infinitives unlike *for*, because the affix -*enne* was a morphological realization of dative Case. In covert syntax, [−wh] *c*

attracts formal features of *-enne* including Case:

(37) $[_{cP}\ c(-\text{wh})\ [_{CP}\ C\ [_{VP}\ \text{V-enne}\ ]]]$

Case

To check off this Case feature, the complementizer must include the upper *c*, and consequently the amalgamate is always spelled out as *to* at PF.[9]

5.4.2. The Shift of the [−finite] Feature

Having established the structure of *to* infinitives in OE, let us consider how the structure in (36) changed during ME and early Modern English. A first change was the emergence of *for to* infinitives in early ME. As Visser (1963-1973) and Nakao (1972) argue, *for* was introduced in the late 11th century to reinforce the directional and purposive meanings of *to*. Since these meanings are closely related to the clause type indicator *c* as discussed above, the introduction of *for* suggests that *to* gradually lost its original directional meaning and came to serve as a pure subordinator C, and *for* took over the former function of *to* by realizing the upper *c*. Thus, the structures of *to* infinitives and *for to* infinitives in early ME can be delineated as follows:

(38) *Early ME*

a. $[_{cP}\ c(-\text{wh})\ [_{CP}\ C\ [_{VP}\ \text{V-enne}\ ]]]$ (*to* infinitives)

⇩

to

b. $[_{cP}\ \text{for}\ [_{CP}\ \text{to}\ [_{VP}\ \text{V-enne}\ ]]]$ (*for to* infinitives)

In this period, old type *to* infinitives where *to* was a complex category and new type *for to* infinitives where *for* and *to* occupied distinct heads of the split-CP structure coexisted. In either case, the dative Case of the affix *-enne* is checked off by the upper *c* in the same way as in OE. It is also important to note that at the present stage, the *for* NP *to* VP construction is not permissible, since [Spec, CP], an A-bar topic position, is not available for a lexical subject of infinitives. In light of general prohibition against topicalization within infinitives, there is no motivation for raising a lexical subject from the VP-internal original position to [Spec, CP].

It may be worth pointing out that just as *c*+C was spelled out as a single word *to* in *to* infinitives, *for to* was often spelled as one word:

(39) ȝif eni mon bit *fort* iseon ou
if any man asks to see you

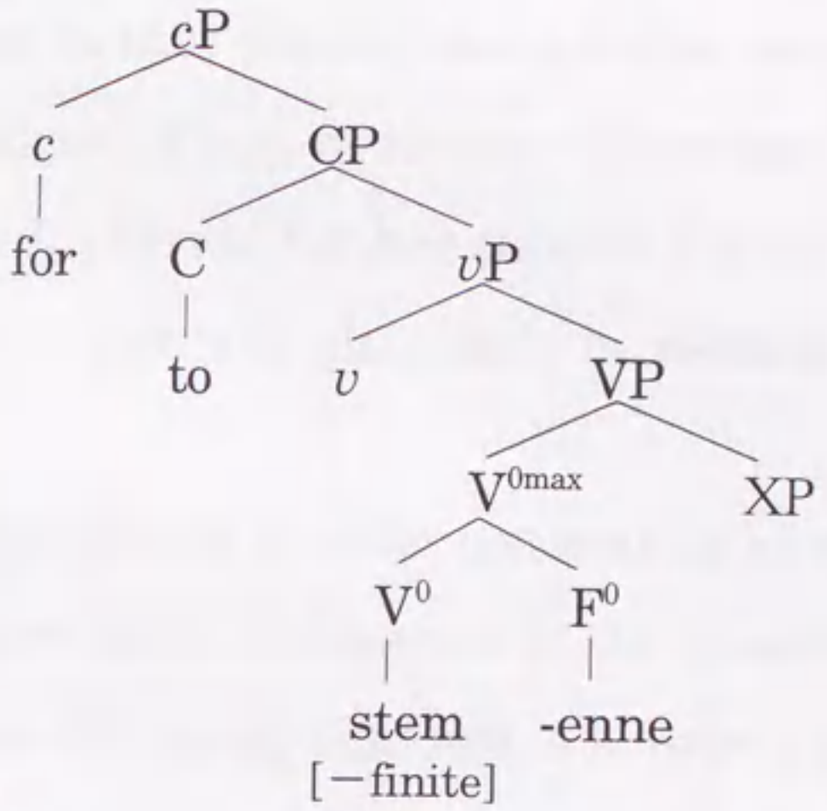
(*The Ancrene Riwle* 41, Mustanoja (1960:515))

I have claimed that *c* and C undergo a morphological Merger operation on their way to PF and the resulting complex category is pronounced as a single word. The fact that *for* and *to* often amalgamated into one word in early *for to* infinitives serves as evidence showing that this operation is actually at work in the phonological component.[10]

A second change that took place in ME infinitives was the demise of the infinitival affix *-enne*. It slowly declined through ME and came to be spelled as *-en* or *-e*. Along with this decline, inflected

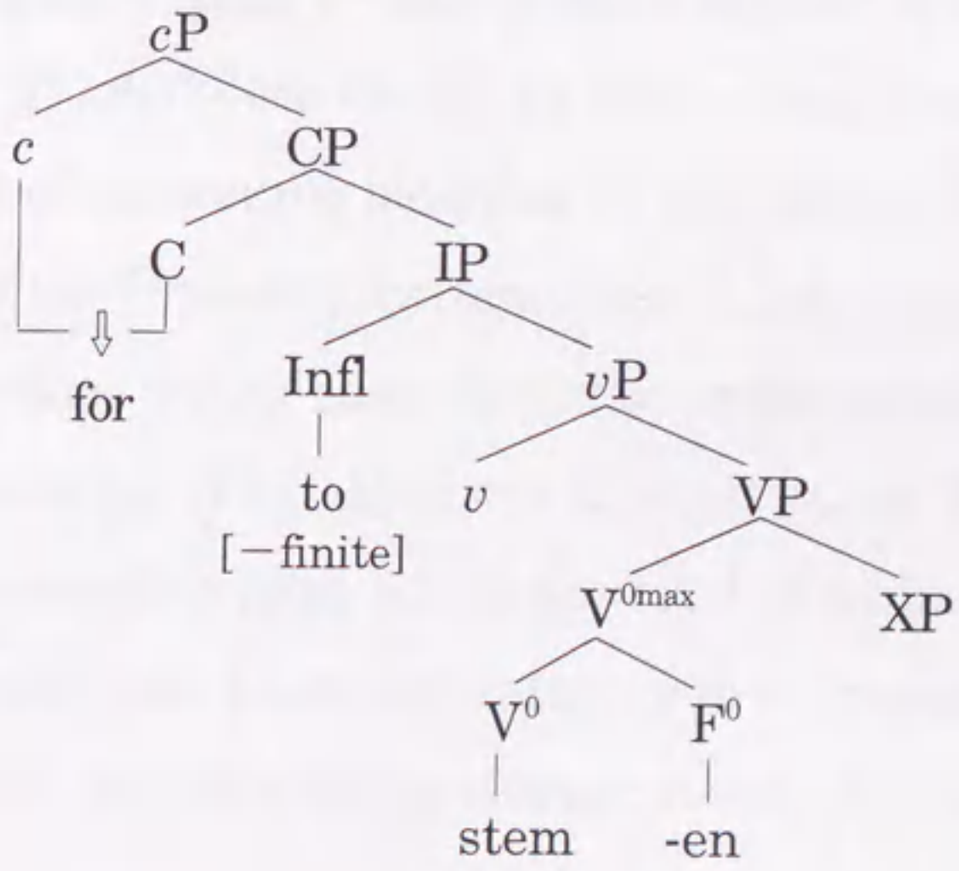
infinitives gradually lost its nominal properties and came to be interpreted as pure verbal projections. Then, I propose that this change brought about the following two effects: (i) the dative Case of the affix became optional, and (ii) the affix lost its ability to carry the [−finite] feature, and it came to be realized by another element, namely, *to*. This means that *to* was reanalyzed as a tense element and the Infl node was introduced into infinitival clauses. The change can be represented as follows:

(40) a.



(Early ME)

b.



(Late ME)

As the tree diagrams indicate, two factors combined to cause the reanalysis: the downward shift of *to* from C to Infl and the upward shift of the [−finite] feature from F^0 to Infl. Since *to* had already lost its directional meaning and served as a pure subordinator in early ME, it was available as the host of the [−finite] feature, which could no longer be carried by the infinitival affix. Compared to the shift of [+F] in finite clauses discussed in Chapters 3 and 4, the diachronic feature shift in infinitival clauses is somewhat complicated, but they are quite similar in that the decline of verbal morphology causes parametric changes. Thus, as long as the present analysis is on the right track, it will provide further support for our hypothesis that parameter resetting should ultimately be reduced to morphological changes.

Here, a comment is in order concerning locality property of diachronic feature shift. In Chapter 3, the generalization was drawn that in a process of language change, a feature on a functional head X^0 can shift to another head Y^0 only if there is no Z^0 intervening between X^0 and Y^0 (cf. (34) in Ch.3). In light of this generalization, it is predicted that *v* intervening between F^0 and Infl in (40a, b) would block the shift of the [−finite] feature. Recall, however, that children fix parametric values on the basis of surface word order patterns of primary linguistic data (PLD) they are exposed to, not base-generated structures like (40a, b). At the level of surface patterns, *v* and V are combined into a complex category to be pronounced as a single word. Thus, the parametric change in infinitival clauses is

readily described as the shift of the [−finite] feature from F^0 contained in the verbal complex Vb to the newly introduced Infl node, as in (41):

(41) $[_{cP}\ c\ [_{CP}$ to $[_{vP}$ Vb $[_{VP}\ t_V\ ]]]]$ (Early ME)
[−finite]

⇩

$[_{cP}\ c\ [_{CP}\ C\ [_{IP}$ to $[_{vP}$ Vb $[_{VP}\ t_V\ ]]]]]$ (Late ME)
[−finite]

Since the verbal complex Vb and the infinitive marker *to* are adjacent at the level of the surface word order patterns in (41), the shift of the [−finite] feature can be executed without any acquisitional burden on the part of children. It thus can be said that the parameter resetting under discussion is consistent with the general locality property of diachronic feature shift.

Given this reanalysis, the structures of *for to* infinitives and *to* infinitives in late ME are delineated as follows:[11]

(42) *Late ME*

a. $[_{cP}\ c$(−wh) $[_{CP}\ C\ [_{IP}$ to $[_{VP}$ V-en $]]]]$
⇩
for
(*for to* infinitives)

b. $[_{CP}\ C\ [_{IP}$ to $[_{VP}$ V-en $]]]$ (*to* infinitives)

Along with the shift of *to*, *for* was also reanalyzed as a complex category consisting of *c* and C, so that it took over the former function of *to* completely. The structure of *for to* infinitives in late ME in (42a) is different from that in early ME in (38b), in which *for* and *to* realized each distinct head of the split-CP structure. Note

that since dative Case of the infinitival affix became optional, [−wh] *c* serving as a Case-checker was no longer obligatory. When simple CP lacking *c*P appears, a *to* infinitive without *for* is obtained, as in (42b). Moreover, as we have reviewed in section 5.2, the *for* NP *to* VP construction emerged in this period. Due to the introduction of the Infl node, [Spec, IP] became available for the lexical subject of infinitives:

(43) a. $[_{cP}$ *c*(−wh) $[_{CP}$ C $[_{IP}$ Subj to $[_{VP}$ V-en]]]]
Case (*for* NP *to* VP)

b. $[_{cP}$ *c*(−wh) $[_{CP}$ C $[_{IP}$ to $[_{VP}$ V-en]]]]
Case (*for to* infinitives)

In OE and early ME, the Case-assigning feature of the upper *c* was always checked off by the corresponding Case feature of the infinitival affix *-enne*. Once dative Case of the affix became optional, *c* came to be able to act as a Case-checker for the lexical subject as in (43a). If the affix *-en* has a Case feature, on the other hand, it must be attracted by *c* in covert syntax so that no lexical subject can appear in [Spec, IP], as shown in (43b). In this way, *for to* infinitives and the *for* NP *to* VP construction coexisted in this period (see the chronological table (21)). If there is no lexical subject nor does the affix *-en* have a Case feature, *c* is not necessary and thus a simple *to* infinitive in (42b) is derived.

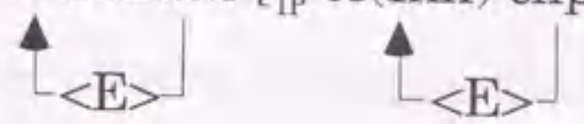
For the explanation given above to be valid, we must verify the hypothesis that the Infl node was introduced into infinitival clauses in late ME. An effect of the introduction of Infl manifested itself in

causative verb complements. It was argued that in OE, both direct and indirect causatives were expressed by bare infinitival complements. However, they were expressed by different ways in late ME, as pointed out by Fischer (1995). Consider the following example:

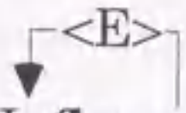
(44) And slepynge in hir barm upon a day,/ She *made to clippe* or shere his heres away,/ And *made* his foomen al his craft *espyen*.
'And one day, while he was sleeping in her bosom,
she made (someone) cut off his hair and let the enemy
see his secret power.'
(Chaucer *Monk* 2064-2067, Fischer (1995:1))

In this sentence, causative *make* with a *to* infinitive and that with a bare infinitive are coordinated, but they have different interpretations. As for the case of *made to clippe*, Fischer quotes a Latin translation that says *vocavitque tonsorem et rasit* 'and [she] called a barber and [he] cut' and observes that the causal relation is indirect. By contrast, the case of *made. . . espyren* is naturally interpreted as a direct causation: it reads as 'she enables the enemy to see his secret power by intentionally unveiling it'. From this example and on many other reasonable grounds, Fischer concludes that bare infinitives in late ME were used exclusively to express direct causation. Using the notation with <E> features, we can represent the contrast as follows:

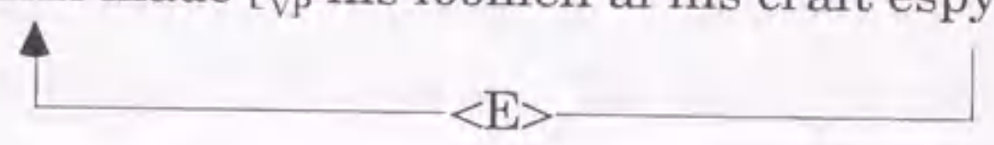
(45) a. [$_{IP}$ she Infl made [$_{IP}$ to(Infl) clippe his heres away]]



(indirect causation)



b. [$_{IP}$ she Infl made [$_{VP}$ his foomen al his craft espyen]]



(direct causation)

If Fischer's observation is correct, it suggests that the infinitival affix *-en* was no longer able to carry the [−finite] feature unlike OE *-an* and *-enne*. The only tense element available for the <E> feature of *espyen* in (45b) is the matrix Infl, where it amalgamates with the <E> feature of the causative predicate. On the other hand, when a *to* infinitive is used as in (45a), *to* acts as a host of the [−finite] feature, so that this configuration yields the indirect causative interpretation.

Another empirical phenomenon that serves as evidence for the introduction of Infl node is the rise of passive infinitives in late ME. After the verbs *do* and *make*, *to* infinitives were the rule when they were in passive forms (Fischer (1995:15)), as in (46):

(46) . . . he *dyde* Dabeney *to be arestyd* for mayntening . . .
he made Dabeney to be arrested for maintaining
(*The Paston Letters* 181.19-20, ibid.)

This indicates the EPP effect showed up within infinitives. To obtain the word order in the embedded clause, [Spec, IP] must be available for the derived surface subject *Dabeney* and the head Infl must have a strong D-feature that induces overt movement of the underlying

object to the Spec position. Thus, the connection between passive forms and a preference for *to* infinitives in late ME suggests that *to* was recategorized as a member of Infl and causative verbs became able to take IP complements.[12]

Although the infinitival affix -*en* lost the ability to carry the [−finite] feature in late ME, it is still important to bear in mind that the affix could optionally assume a Case feature in late ME. Evidence that *to* infinitives behaved like nominals can be seen in the following examples:

(47) a. Ne i herd neuer . . . in land Men sua hard *at*
not I heard never in land men so hard at
to vnderstand.
understanding

(14c. *Cursor Mundi* 9326, Visser (1963-1973:1032))

b. Vyenne salewed parys *wythoute to make* ony
Vyenne saluted Parys without making any
semblaunce of Love.
semblance of love

(c 1485 *Parys & the Fair Vyenne* 1.32, ibid.:1035)

c. *In stede to healpe* hym to dye well, [he] putteth hym in
instead of helping him to die well he puts him in
vayne hope of long lyfe.
vein hope of long life

(1557 North, tr. Gueuara's Diall Pr. 231 a/1, ibid.:1033)

To infinitives cooccur with a wide variety of prepositions other than the infinitival complementizer *for*. In late ME, this construction was less productive than in early ME, but it was permitted until early

Modern English. The prepositions above clearly have Case-assigning features which must be checked off by some way or other. Thus, this construction is allowed only when the infinitive has a corresponding Case feature that is attracted by the preposition in covert syntax.

5.4.3. The Demise of For To *Infinitives*

Even after the rise of the *for* NP *to* VP construction, *for to* infinitives survived until the mid-16th century. An important point to note is that the infinitival affix *-en* was also finally lost in the same period (cf. Lightfoot (1979) and Roberts (1993)). To account for the simultaneous demise of *for to* infinitives and the affix *-en*, let us consider the realization of external arguments in infinitives. Tanaka (1994) argues that the affix *-en* had the ability to realize the external argument of embedded predicates on the basis of the observation that the so-called *faire-par* (FP) construction in (48), in which external arguments of causative verb complements are not overtly expressed, was permissible until early Modern English where the affix *-en* still survived:

(48) Than [he] *did* *to bringe* ston and morter.
then he made to bring stone and mortar
(*Merlin* 27, Visser (1963-1973:1347))

In present-day English, this construction is ruled out since PRO cannot be properly licensed in the subject position of causative verb complements. Drawing the generalization that only languages with infinitival affixes allow the FP construction, Tanaka claims that the

external argument of the infinitive in (48) is not PRO, but realized as the morpheme *-e*. As expected, this construction was lost in the 16th century along with the final demise of the infinitival affix.

Essentially adopting Tanaka's proposal, let us assume the following property of the infinitival affix:

(49) The infinitival affix optionally functions as an external argument. (cf. Tanaka (1994))

On this assumption, the structures of *for to* infinitives before and after the mid-16th century will be represented as in (50):

(50) a. $[_{cP}$ *c*(−wh) $[_{CP}$ C $[_{IP}$ to $[_{VP}$ V-en]]]]

Case (before mid-16c)

b. *$[_{cP}$ *c*(−wh) $[_{CP}$ C $[_{IP}$ PRO to $[_{VP}$ V]]]]

No Case (after mid-16c)

Under the present framework, *for to* infinitives necessarily include *c*, because the complementizer *for* is a morphological realization of *c* (early ME) or *c*+C (late ME). In either case, the Case-assigning feature of *c* must be checked off. Before the mid-16th century, the infinitival affix *-en* functioned as the external argument if there is no lexical subject, and the Case-assigning feature of *c* could be checked off by the dative Case of *-en*. On the other hand, once the affix was lost in the mid-16th century, the external argument of infinitives came to be realized as PRO.[13] Since PRO has no Case feature to be checked off against *for*, the Case-assigning feature of *c* in (50b) will remain unchecked, causing the derivation to crash. Therefore, a

lexical subject became obligatory whenever *for* appeared in infinitival clauses and consequently *for to* infinitives died out.

To recapitulate the history of *for to* infinitives in English, the relevant factors can be summarized as follows:

(51)

	1000	1100	1200	1300	1400	1500	1600
for			*c*		*c*+C		*c*+C
to	*c*+C		*c*+C or C		Infl		Infl
inflection	-enne		-enne		-en/-e		ϕ
dative Case	obl.		obligatory		optional		absent
	(A)		(B)		(C)		(D)

This shows that the functional change of the complementizers took place hand in hand with the loss of the infinitival affix *-enne*. When *for* was introduced as a new complementizer, *to* came to function as a pure subordinator C, and it was finally reanalyzed as an Infl element. At the same time, as the affix *-enne* declined to be spelled as *-en* or *-e*, it lost the ability to carry the [−finite] feature, so that this feature shifted to the newly introduced Infl element, *to*.

In the period (A), there were no *for to* infinitives because *to* was still active and retained its directional and purposive meanings. Once *for* was employed as the clause type indicator *c*, however, *to* was reinterpreted as C. This brought about the rise of *for to* infinitives in the period (B). Later in the period (C), *to* further changed into an Infl element bearing the [−finite] feature, and the

Case feature of the affix *-en* became optional. This made [Spec, IP] available for the lexical subject and gave rise to the emergence of the *for* NP *to* VP construction. Finally in the period (D), when the affix was completely lost and PRO appeared in infinitives, *for to* infinitives became obsolete because the Case-assigning feature of *for* was no longer properly checked off by a non-lexical external argument.

5.5. The Spread of the ACI Construction

Since the structural development of *to* infinitives has been made clear, we can turn to the spread of the ACI construction to B-type verbs in late ME. As we have seen in section 5.3.1, this construction emerged in the mid-14th century, approximately the same period as the rise of the *for* NP *to* VP construction. It will be shown below that their simultaneous development was not a mere coincidence but induced by the same factor, that is, the introduction of the Infl node into infinitival clauses.

5.5.1. S-Selection and Functional Categories

To investigate the relevance of Infl to the ACI construction with B-type verbs, let us focus our attention on the s(emantic)-selectional property of the matrix predicates. It is generally true that complements of B-type verbs express some propositional content; in other words, the semantic property common to all B-type verbs is that they s-select [proposition]:

(52) believe:_____[proposition]

The term "proposition" is defined here as a semantic unit that has its own truth value. Consider (53):

(53) a. She believes that proposition.

b. She proved that fact.

c. *She believes that state of affairs.

d. *She proved this eventuality. (Bach (1977:641))

Noun phrases that denote some event or state of affairs without truth values cannot appear as complements of B-type verbs. The (un)grammaticality of the examples above is expected from the selectional restriction in (52). On the other hand, clausal complements of perception verbs do not express propositional contents. The following contrast concerning the occurrence of modal adverbs proves this point:

(54) a. Bill thinks that Carol probably lied.

b. *Bill saw Carol probably running in the park.

In (54a), the modal adverb *probably* can occur within the clausal complement embedded under the B-type verb *think*. This adverb modifies the content of the proposition (cf. Nakano (1996)). By contrast, the same adverb cannot occur in the perception verb complement in (54b), which shows that the embedded clause does not have any propositional content.[14] It should also be added that perception verb complements are always included in the scope of negation in the matrix clause, whereas such implicational relations do not exist between B-type verbs and their complements:

(55) a. We didn't see him leave.

⊃ He didn't leave.

b. John didn't believe that Mary was ill.

⊅ Mary was not ill.

c. She doesn't believe John to be honest.

⊅ John is not honest.

(a, b from Nakano (1999:22))

In the perception verb construction, whenever the state of affairs expressed by the matrix clause is true, the state of affairs in the embedded clause must also be true; on the other hand, because B-type verb complements have their own truth values, the content of the embedded clause may not be true even if the state of affairs in the matrix clause is true. It thus can be said that unlike perception verb complements, B-type verb complements express propositional meanings.[15]

It must be noted that when *see* means "find out, understand," its complement can express propositions, but in that case, the complement is always realized as a finite clause or an IP infinitive:

(56) a. At first I did not see that she was so stingy.

b. We saw the project to be impracticable.

If we assume with Safir (1993) and others that perception verb complements are VPs, then we are led to conclude that a VP complement cannot express a proposition with its own truth value. To put it intuitively, although bare VP infinitives can express events or state affairs, they must be further dominated by some functional

category to express propositions. From these grounds, I propose the following condition:

(57) [Proposition] is a semantic feature that is carried by some functional category.

Let us assume that the [proposition] feature can only be carried by a "pure" functional head that is specified as [−lexical, +functional] so that *v*, which is specified as [+lexical, +functional], is not qualified as a host of this feature (see note 1 in Ch.2). If this assumption is on the right track, we can straightforwardly explain why B-type mental perception verbs cannot select bare infinitival complements. Given that bare infinitival clauses are VPs (or *v*Ps), they are not able to carry the [proposition] feature. Thus, the functional category Infl in English infinitives plays a crucial role in expressing propositional meanings in the ACI construction with B-type verbs.

We essentially adopt here a selectional approach to θ-roles. This is to say that [proposition] is not assigned to the whole IP or *c*P by the matrix predicate, but it is an independent semantic feature that is selected by the matrix predicate. Evidence for the existence of the semantic feature [proposition] is provided by the well-known restriction on ECM infinitives: they cannot cooccur with event predicates, as exemplified in the following examples:

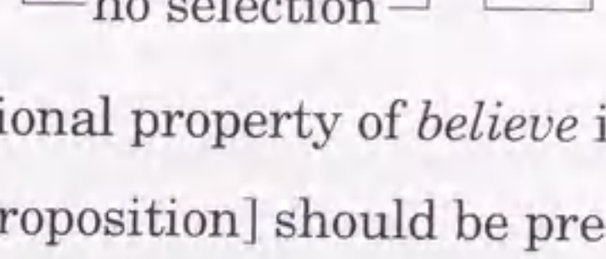
(58) a.*John believed Mary to bring the beer.

b.*I proved John to bring the beer. (Martin (1992:15))

The event feature <E> is again relevant here. Suppose that Infl has only one slot that determines the propositional/eventive

interpretation of infinitives. This assumption captures the insight that single syntactic units such as IP cannot have both propositional and eventive interpretations at a time. Under the proposed system, then, the [proposition] feature would occupy the same slot to which <E> has been discharged:

(59) ··· believed [$_{IP}$ Mary to(Infl) bring the beer]



In (59), the s-selectional property of *believe* is not satisfied because the slot in which [proposition] should be present is already occupied by <E> that is discharged from *bring*. It thus can be said that the feature conflict between <E> and [proposition] prevents B-type verbs from occurring with event predicates, which in turn verifies the existence of [proposition]. It may be worth mentioning, however, that B-type verbs can take event predicates as their complements when they are embedded under auxiliary *have*, as in (60a), or express habitual meanings, as in (60b):

(60) a. John believed her to have come here.

b. John believed the cat to catch a mouse.

This seems to be a potential counterargument against our approach, but the present analysis can deal with these examples as well. As for (60a), auxiliary *have* is an overt realization of the embedded reference time (cf. Hornstein (1990)), so that it can receive <E> discharged from the past participle:

(61) ⋯ believed [$_{IP}$ her$_i$ to(Infl) [$_{FP}$ have [$_{VP}$ t_i come here]]]

[proposition] [<E>]

selection

The selectional property of *believe* is also satisfied because it can properly select [proposition] located in the embedded Infl. (60a) is thus acceptable. As for (60b), the event specification of a habitual verb does not have to be anchored to a specific reference time, so that the <E> feature of *catch* need not be discharged to a tense element (see (34)). Again, the embedded Infl carries [proposition] that is selected by *believe* and hence the sentence is acceptable. The grammaticality of (60b) is thus explained on a par with familiar ACI sentences like *John believes Mary to be innocent*, where no eventive interpretation is involved in the embedded infinitival clause.

Since the condition in (57) does not specify any particular functional category that contains the [proposition] feature, the wide range of structural realizations of the complement of B-type verbs is naturally accounted for. When [proposition] is located in D, we get noun phrase complements as in (53), repeated here:

(62) a. She believes [$_{DP}$ that proposition].

b. She proves [$_{DP}$ that fact].

The ungrammaticality of (63a) can be explained in much the same way as we did for (58):

(63) a.?*We were aware of the decision to leave the company.

(Ormazabal (1994:479))

b. ··· be aware of [$_{DP}$ PRO the(D) decision to leave ···]

[<E>]

no selection

Be aware of, like *believe*, s-selects [proposition] and cannot occur with event nominals such as *decision*. In (63b), the slot where [proposition] should be present is already filled with <E> discharged from *decision*; consequently, the selectional property of *be aware of* is not satisfied when its complement describes an event, not a proposition as in (62).

When *c* is selected as the locus of [proposition], the complement of B-type verbs is realized as *that*-clauses. What is interesting here is that whereas ECM infinitives do not cooccur with event predicates, *that*-clauses escape from this restriction:

(64) a. *I believe Mary to arrive tomorrow.

b. I believe that Mary arrives tomorrow.

The grammaticality of (64b) is correctly predicted under the present approach:

(65) ···believe [$_{cP}$ that(*c*) [$_{CP}$ t_C [$_{IP}$ Mary Infl arrives ···]]]

[proposition] [<E>]

selection

Since [proposition] is in *c*, it does not conflict with <E> discharged to Infl by *arrive*; thus, the s-selectional restriction on the complement of *believe* is properly satisfied.

A question then arises as to how small clause complements of B-type verbs should be dealt with in light of the condition in (57). If

small clauses are maximal projections of N or A, they would not be qualified to express propositions for the same reason that bare VP infinitival complements cannot. Chomsky (1993) analyzes small clauses as AgrP on the basis of the split-Infl hypothesis of Pollock (1989), but later in Chomsky (1995), he abandons this hypothesis and eliminates Agr altogether; as a result, small clauses are assumed to be projections of embedded predicates. Nevertheless, there is empirical evidence indicating that some functional category is included in small clauses: the effect of the EPP is observed within small clauses:

(66) a. I found [it impossible to persuade him].

b. *I found [ϕ impossible to persuade him].

When an embedded predicate has no external argument, the expletive *it* is obligatorily inserted into the Spec of the small clause. Given that the strong D feature that induces the effect of the EPP cannot be carried by substantive categories (Chomsky (1995:232)), expletive insertion in (66) strongly suggests that small clause complements of B-type verbs involve a functional category. Let us suppose, then, that small clauses are IPs, following Hornstein and Lightfoot (1987) and Aarts (1992). The relevant structure can be delineated as follows:

(67) I believe [$_{IP}$ John$_i$ Infl [$_{AP}$ t_i innocent]]

[proposition]

selection

In this case, [proposition] is located in Infl and selected by the matrix

verb. Aarts (1992) posits a structure in which VP headed by a phonetically null verb dominates AP or NP within small clauses, but I would rather assume that AP or NP is directly generated as the complement of Infl, as in (67). Although this Infl node is inert with respect to tense specifications, the agreement relation between the embedded predicate and its subject is overtly marked in languages like French. Even in English, number agreement relations hold within small clauses in such examples as *I consider these teachers megalomaniacs/*a megalomaniac* (Aarts (1992:181)). Moreover, the present analysis in fact does not contradict with the leading idea of Chomsky (1995). He proposes that only functional categories that have features interpretable at LF are legitimate syntactic objects. Infl in small clauses is considered to be legitimate because it carries the [proposition] feature which receives an interpretations at LF.

5.5.2. An Explanation for the Historical Change

Now that the mechanism of s-selection of B-type verbs has been made clear, we are ready to consider the development of the ACI construction in the history of English. As we have seen in section 5.3.1, the ACI was allowed only with causative and perception verbs in OE, and B-type verbs could take *that*-clauses, small clauses, and control infinitives as their complements. The structures of the OE examples in (16) and (18) can be represented as follows:

(68) *Causative/Perception Verb Complements*

a. se . . . doeð [$_{VP}$ hie cwacian]
that makes them quake-Inf

b. Ic seah [$_{VP}$ turf tredan]
I saw turf tread-Inf (=(16))

(69) *B-Type Verb Complements*

a. he secge oððe gelyfe [$_{cP}$ ðæt þry Godas syndon]
he say or believe that three Gods are

b. þa þing [Op$_i$ þe ge . . . [$_{IP}$ t_i Infl betstan] gelyfað]
the things that you best believe

c. Ic gelyfe [$_{cP}$ to [$_{CP}$ t_C [$_{VP}$ arisenne on domes dæge]]]
I believe to rise on doom's day (=(18))

The causative/perception verb complements in (68) raise no problems, since they have only eventive interpretations so that the matrix verbs can select VP complements without violating the condition in (57), which requires the presence of a functional category for a propositional interpretation. Complements of the B-type verb in (69), on the other hand, all include functional categories. The [proposition] feature is contained in *c* in the *ðæt*-clause complement and Infl in the small clause complement. What concerns us most is the control infinitive in (69c). Here, *c* is selected as the locus of [proposition] because OE infinitives lack an independent Infl node. The head *c*, together with the subordinator C, is morphologically realized as *to*.

It is important to note that besides the complementation patterns given in (69), B-type verbs in OE occasionally took bare

infinitival ACI complements. Consider the examples below:

(70) a. se æt Heorote fand wæccendne wer wiges bidan
who at Heorot found walking man battle abide-Inf
'who found the walking man waiting for battle at Heorot' (*Beowulf* 1267, Fischer (1989:192))

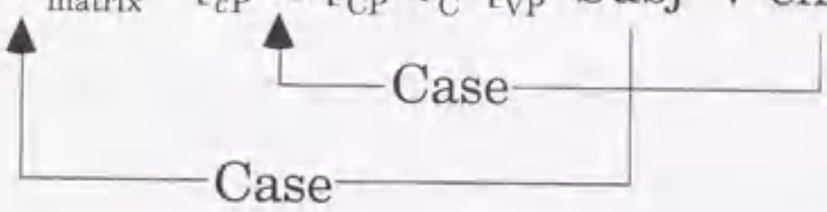
b. Ic wat eardfæstne anne standan, deafne, dumban,
I know firm one stand-Inf deaf dumb
'I see someone stand firm, deaf, dumb . . . '
(*Riddles* 49.1, ibid.:201)

B-type verbs like *find* and *know* in the present-day English take ACI complements with propositional meanings. If the bare infinitival complements above also have propositional interpretations, they will constitute counterexamples to the condition in (57). However, Fischer (1989) observes that OE *findan* and *witan* can be followed by a bare infinitives only if the infinitive conveys some physical, concrete action that is simultaneous with the action expressed by the matrix verb. If this observation is correct, then *findan* and *witan* should be classified as perception verbs when they take bare infinitival ACI complements. It thus can be said that these examples do not pose any problems for our analysis; rather, they serve as evidence that clausal complements that lack functional categories cannot express propositional meanings.

We are now in a position to offer an explanation for the development of the ACI construction. It must be made clear (i) why the ACI with *to* infinitives was not allowed in OE and early ME, (ii) why this construction became permissible in late ME, and (iii) why

control infinitives became obsolete after B-type verbs along with the introduction of the ACI. The first point can be accounted for in terms of the Case-licensing of the embedded subject. If a B-type verb in OE took an ACI complement, the relevant structure would be delineated as follows:

(71) *V_{matrix} [$_{cP}$ to [$_{CP}$ t_C [$_{VP}$ Subj V-enne]]]



The matrix verb must select *c*P when the infinitive conveys some propositional content due to the condition in (57). As argued in section 5.4.1, the complementizer *to* checks off the dative Case feature of the infinitival affix *-enne*. If a lexical subject is additionally involved in the embedded clause, it must be Case-checked by the matrix verb in covert syntax. An important point to note is that formal features of the subject necessarily cross the *c*P boundary when they move to the matrix verb. Under the framework of Chomsky (1995), covert feature raising of a nominal phrase counts as A-movement (p.272). In view of the general prohibition against A-movement across a *c*P boundary, it is reasonably concluded that the derivation in (71) should be ruled out.[16] Another possibility is that the embedded subject has dative Case which would be checked off by the complementizer *to*. In this case, the affix *-enne* would have to be Case-checked by the matrix verb. This derivation, however, is also properly ruled out, since the matrix verb only has the ability to check accusative Case and does not license the dative Case of the

affix -*enne*. Therefore, a lexical subject cannot appear in the infinitival complement of B-type verbs in either case.

Once the Infl node emerged in infinitival clauses in late ME, this newly introduced node came to serve as the locus of the [proposition] feature. Furthermore, since the dative Case of the infinitival affix -*en* became optional in late ME, *c* was no longer necessary in infinitives when the affix did not have Case. It follows that B-type verbs came to select an IP complement:

(72) V_{matrix} $[_{IP}$ Subj to $[_{VP}$ V-en]]
Case

The Case feature of the embedded subject must be checked off by the matrix verb. Since no *c*P boundary intervenes between the matrix verb and the embedded subject, covert feature raising for Case-checking is not blocked, contrary to OE and early ME. In this way, the ACI construction spread to B-type verbs in late ME along with the reanalysis of *to* from a complementizer to an Infl element.[17]

Finally, let us consider the demise of control infinitives with B-type verbs. The realization of external arguments is again relevant here. In section 5.4.3, we adopted Tanaka's (1994) proposal that the infinitival affix optionally functions as an external argument. To put it more precisely, he claims that the infinitival affix functions as an argument if and only if it is Case-marked (see note 13). Let us assume this to be the case. If a B-type verb in late ME selects the [proposition] feature in Infl and takes an IP control infinitive as its complement, the embedded external argument would

be realized as PRO, since the infinitival affix cannot be Case-marked due to the absence of *c*. Thus, the structure would be as follows:

(73) V_{matrix} $[_{IP}$ PRO to $[_{VP}$ V-en]]

This structure can be straightforwardly ruled out because PRO in [Spec, IP] is not properly licensed. Under the traditional binding-theoretic approach to PRO, the matrix verb in (73) governs PRO, which leads to a violation of the PRO theorem. An alternative Case-theoretic approach to PRO advanced by Martin (1992) also rules out this structure. On the assumption that PRO has a null Case, he distinguishes two types of non-finite Infl nodes in terms of the feature [±tense] and associates its values with the ability to check a null Case. Specifically, he proposes that only the non-finite Infl node with the value [+tense] can check a null Case. Here, [+tense] means that the Infl node has a time frame that is not simultaneous with the matrix clause and represents an unrealized event, whereas [−tense] Infl does not have its own time frame so that it is interpreted as totally dependent on the matrix tense. Since ECM infinitives in present-day English cannot represent unrealized events, it can be said that B-type verbs select [−tense] Infl when they take infinitival IP complements.[18] It follows, therefore, that PRO in (73) cannot be Case-checked, causing the derivation to crash.

5.5.3. Crosslinguistic Evidence

From our explanation for the historical development of the ACI

construction, it is predicted that only languages that have an independent Infl node within infinitives allow the ACI with B-type verbs. In this section, we investigate French and Icelandic, and demonstrate that this prediction is really borne out.

5.5.3.1. French

It is a well-known fact that French B-type verbs can take control infinitives but cannot take ACI infinitives as their complements, as shown in (74):

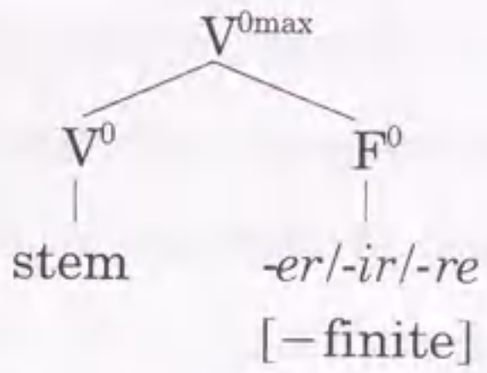
(74) a. Je crois [PRO avoir fait une erreur].
I believe have-Inf made a mistake
'I believe that I have made a mistake.'

b. *Jean croit [Bill avoir menti].
John believes Bill have-Inf lied
'John believes Bill to have lied.' (Kayne (1984:103, 112))

Besides control infinitives, *croir* 'believe' can be followed by small clauses and finite clauses. Thus, it exhibits the same complementation patterns as those of OE B-type verbs.

French infinitives have inflectional affixes *-er/-ir/-re* which distinguish them from finite forms. Under the present framework, they are expected to have the following word internal structure:

(75) *French Infinitives*



Like OE infinitives, the inflectional affixes are able to bear the [−finite] feature. In addition, given that French infinitive marker *de* is a complementizer (see (28a)), there seems to be no need to posit an independent Infl node within French infinitives in view of economy considerations for syntactic representations.

The absence of Infl in French infinitives is supported by the lack of the EPP effect. A striking difference between English and French infinitives is that in English raising verbs can appear as the complement of causative verbs whereas in French they cannot:

(76) a. His pained expression makes John seem to be suffering.

b. *Son expression peinée fait sembler Jean
his expression pained makes seem-Inf John
souffrir.
suffer-Inf (Kayne (1975:254))

The ungrammatical French example in (76b) can be ruled out by means of the Predication Condition proposed by Rothstein:

(77) *The Predication Condition*

Every syntactic predicate must be syntactically saturated. (cf. Rothstein (1995:503))

Suppose that in order for a predicate to be syntactically saturated, it

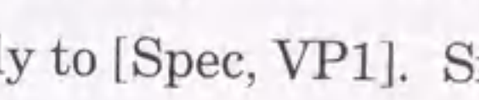
must be c-commanded by an appropriate DP at S-structure or before Spell-Out. Raising verbs such as *seem*, which do not have external arguments, must be c-commanded by an expletive or DP that has raised out of the complement clause.[19] In (76b), *Jean* remains within the infinitival complement of *sembler*, resulting in a violation of the Predication Condition. Then, the issue is why the operation that raises *Jean* cannot be applied in (76b), as in (76a), to save the violation. I claim that this is due to the difference in the categorial status of infinitival complements in these languages. Here, we will follow the Minimalist assumption that an operation applied to satisfy the Predication Condition must be independently motivated by a formal requirement to check some feature. In other words, the condition (77) itself is not a trigger for movement. With this assumption in mind, let us first look at the derivation of the grammatical English example. If English infinitives are IPs as is widely assumed, the derivation proceeds as follows:

(78) $[_{IP1}$ John$_i$ $[_{VP1}$ seem $[_{IP2}$ t_i to $[_{VP2}$ t_i be suffering]]]]

The subject *John* moves from the most deeply embedded [Spec, VP2] to [Spec, IP1] to satisfy the EPP. Chomsky (1995) observes that the EPP effect is reduced to a strong D-feature of Infl. If this is the case, then the movement of *John* must be overt: otherwise the derivation would crash, because a strong feature must be eliminated as soon as it is introduced in a phrase structure (ibid.:234). Since *John* c-commands *seem* before Spell-Out, (78) meets the Predication

Condition. On the other hand, the immobility of *Jean* in (76b) means that there is no EPP effect, which in turn suggests the lack of Infl node in the French infinitival complement:

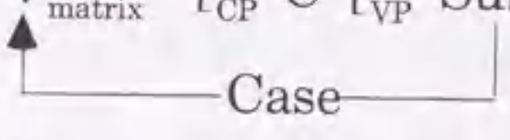
(79) $[_{VP1}$ sembler $[_{VP2}$ Jean souffrir]]



VP1 contains no strong feature to induce the EPP and thus *Jean* does not move overtly to [Spec, VP1]. Since *Jean* does not c-command *sembler* before Spell-Out, (79) violates the Predication Condition.

Let us return to the infinitival complement of B-type verbs. Given that French infinitives lack IP, the [proposition] feature must be carried by some other functional category. Unlike OE, however, the infinitival marker *de* is not morphologically spelled out after B-type verbs in French, which indicates that the full-fledged *c*P-CP structure is not formed. From these grounds, I would propose that French B-type verbs select CP just like English *try*-class verbs:

(80) $*V_{matrix}$ $[_{CP}$ C $[_{VP}$ Subj V-er]]



If a lexical subject appears in the embedded infinitival clause, its formal features must raise to the matrix verb for Case-checking in covert syntax , but this movement is bound to be ruled out since it crosses the CP boundary intervening between the matrix verb and the embedded subject. Therefore, the ACI construction is not allowed with B-type verbs in French and only control infinitives can occur with them.

It must be noted, however, that the acceptability of the ACI

construction improves when the lexical subject undergoes A-bar movement, as in (81):

(81) Quel garçon crois-tu etre le plus intelligent de
which boy believe-you be-Inf the most intelligent of
tous?
all
'Which boy do you believe to be the most intelligent of all?' (Kayne (1984:111))

A crucial point is that although CP always blocks A-movement, it does not block A-bar movement unless its specifier is occupied by an element. The derivation of a sentence involving a [+wh] subject in the infinitival complement proceeds as follows:

(82) quel garçon$_i$ crois$_j$-tu [$_{vP}$ t_i t_j [$_{VP}$ t_j [$_{CP}$ [$_{VP}$ t_i etre le plus intelligent]]]]

Case

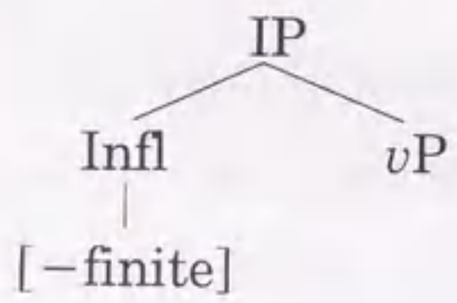
Recall our earlier argument in Chapter 2 that a *wh*-element moves in a successive cyclic way through Specs of *c*P and *v*P (for the definition of Attract F, see (25) in Ch.4). In (82), the *wh*-phrase *quel garçon* overtly moves to the matrix [Spec, *c*P] via the intermediate [Spec, *v*P], leaving copies of formal features in the trace positions. Later in covert syntax, the Case feature of the *wh*-phrase deletes and erases when the matrix verb *crois*, which has raised out of *v*P, attracts the formal features of the *wh*-trace in [Spec, *v*P]. Thus, the A/A-bar asymmetry concerning extraction of the subject from the infinitival complement is reduced to the presence of CP, which is selected by

croire as the locus of the [proposition] feature.

5.5.3.2. Icelandic

Icelandic infinitives uniformly have a weak inflectional affix *-a*. In light of our earlier observation that the weak infinitival affix *-en* in late ME does not have the ability to carry the [−finite] feature, it seems reasonable to assume that Icelandic infinitives include an independent Infl node where [−finite] is located, as in (83):

(83) *Icelandic Infinitives*



This non-finite Infl node serves as a host of the [proposition] feature as well. As expected, Icelandic B-type verbs can take ACI complements:

(84) Ég tel [$_{IP}$ mig hafa á rettu að standa].
I think me-Acc have-Inf of right to stand-Inf
'I believe myself to be right.' (Platzack (1986:218))

The accusative Case feature of the embedded subject *mig* is properly checked off by the matrix verb in covert syntax, since there is no *c*P or CP boundary intervening between them. The acceptability of the ACI with B-type verbs in Icelandic thus naturally follows from the proposed mechanism.

A characteristic of the Icelandic ACI construction is that the embedded infinitival subject can assume inherent Case as well as

structural Case. Consider (85):

(85) a. Jóni batnaði veikin.
John-Dat recovered-from the-disease-Nom
'John recovered from the disease.'

b. Ég tel [Jóni hafa batnað veikin].
I believe John-Dat have-Inf recovered-from the-disease-Nom
'I believe John to have recovered from the disease.'

(Freidin and Sprouse (1991:409))

As for the licensing of inherent Case, let us assume with Chomsky (1981, 1986a) that an argument DP can be inherently Case-marked by the predicate from which it is assigned a particular θ-role. Thus in (85a), Case-marking on the subject *Jóni* (dative) is carried out by the verb *batnaði*. Note here that even when the infinitival form of (85a) is embedded under the matrix verb *tel* in (85b), the subject still assumes dative Case, which indicates that it is Case-marked by the infinitive, not by the matrix verb. One might argue that the embedded clause in (85b) might have French type CP-vP structure with the infinitival subject staying in [Spec, vP], since covert feature raising into the matrix verb for Case-checking is irrelevant in this case. However, there is good reason to suppose that it is also IP: the EPP effect is observed within the infinitival clause under discussion:

(86) Ég tel $[_{IP}$ Jóni$_i$ hafa verið hjálpað t_i].
I believe John-Dat have-Inf been helped
'I believe John to have been helped.' (ibid.:410)

In this example, *Jóni* is inherently Case-marked by the embedded verb *hjálpa ð* at the base-generated position and overtly moves to the surface subject position. Since dative Case is already licensed at the underlying structure, it cannot be a trigger for movement. It can thus be said that the embedded clause has an Infl node with a strong D-feature that induces overt A-movement.

An additional evidence that Icelandic ACI complements are IPs is provided by the fact that infinitives exhibit overt V-to-I movement. In (87), the infinitive *lesa* precedes the VP adverb *alltaf*:

(87) Jón telur [$_{IP}$ í barnaskap sínum marga stúdenta
John believes in foolishness his many students
lesa$_i$ [$_{vP}$ alltaf t_i námsbœkur]].
read-Inf always textbooks
'Lit. John believes many students to always read textbooks in his foolishness.' (Jonas (1996:175))

Given that *alltaf* marks the leftmost edge of *v*P, it is clear from the linear order that *lesa* has raised out of *v*P. Then the subject *marga stúdenta*, which precedes the infinitive *lesa*, must also have overtly raised to Spec of IP, as in (86).

To sum up, the crosslinguistic survey in this section strongly suggests that our historical analysis of the development of the ACI construction is on the right track. French infinitives lack the Infl node and thus ACI complements do not occur with B-type verbs, just like OE and early ME. Icelandic infinitives, on the other hand, are analyzed as IPs, so that B-type verbs can freely take ACI complements in the same way as in late ME onwards. The presence

or absence of Infl within infinitives is determined by the richness of infinitival affixes.

5.6. Summary

In this section, I have offered an explanation for the historical transitions of English infinitival constructions in terms of diachronic feature shift of the [−finite] feature. The main point was that as the infinitival affix *-enne* declined, it lost the ability to carry the [−finite] feature, which brought about the introduction of IP within infinitives. This change took place hand in hand with the reanalysis of *to* from a complementizer to an Infl element. When IP emerged in the mid-14th century, the *for* NP *to* VP construction and the ACI construction with B-type verbs became permissible; and later in the mid-16th century, the infinitival affix was completely lost and consequently *for to* infinitives died out. The relevance of the Infl node to the ACI construction has been confirmed by the survey of French and Icelandic.

Compared to the diachronic shift of [+F] in finite clauses, the parametric change in infinitival clauses has the following characteristics: (i) the shift of the [−finite] feature was upward, and (ii) the Infl node was newly introduced as a result of the feature shift, whereas in finite clauses IP was always present. These differences primarily arise from the fact that only finite clauses have specific tense values: since temporal interpretations of infinitival clauses are given in relation to main clauses, an independent Infl node is not

obligatory. Putting these differences aside, one can immediately notice that the changes in finite and infinitival clauses are quite similar in that both of them are morphologically driven; namely, the decline of inflectional affixes causes parameter resetting. Thus, we can reasonably conclude from the discussion in this chapter that the same kind of mechanism that forces parametric changes in finite clauses is at work in infinitival clauses as well.

notes

* Part of this chapter was presented at the 14th annual meeting of the Modern English Association held at Tohoku Gakuin University on May 23, 1997 and was later published as Nawata (1998b). I am grateful to Hirozo Nakano, Masachiyo Amano, Yoshinobu Niwa, Kozo Kato, Tomoyuki Tanaka, Kazuhisa Ishikawa, Hiroshi Terada, and three anonymous reviewers of *Studies in Modern English* for helpful comments and suggestions on earlier versions of this chapter.

1 Viewed from a historical perspective, the Case-checking property of the complementizer *for* may be attributed to its origin as a preposition. We will return to this issue in section 5.4.

2 In what follows, we will focus our attention on Case-checking by the infinitival complementizer and put aside the problem of whether PRO is Case-checked by the non-finite Infl, because to follow up this matter would involve us in too many factors and carry us far away from the aim of this chapter. For the view that PRO is both formally licensed and semantically interpreted through ϕ-feature checking, see Nawata (1998a).

3 In some dialects including Belfast English and in ME, the complementizer *for* may appear immediately before *to*. In these languages, the possibility cannot be ruled out that *for* and *to* are base-generated under the Infl node as long as *for* does not have any directional or purposive meanings. We will shortly return to *for to* infinitives in ME. For relevant discussion, see also Carroll (1983), Henry

(1995), and Tanaka and Miyashita (1999).

[4] Bošković (1997) claims that *try*-class infinitives are IPs whereas *want*-class invectives are CPs. His approach also structurally distinguishes between *try*-class infinitives and *want*-class infinitives, but the problems concerning the distribution of the null Comp ϕ in (9) and the impossibility of the Subj-*for*-*to*-V word order in (10) will still remain.

[5] Additionally, *try* and *want* can take DP complements. Note that optional Case-assigning features of these verbs need not be checked off when they take clausal complements, since optional Case features are intrinsically inert and do not yield FI violations even if they remain unchecked, as we have seen in section 2.5. This optional Case-assigning feature cannot check off the Case feature of the infinitival subject in cases like *John tried him to leave*, because covert movement of *him* necessarily crosses the CP boundary. We will immediately return to this point.

[6] Postal (1974) classifies the following verbs into *want*-class verbs (i.e. his W-verbs): *expect*, *hate*, *intend*, *like*, *mean*, *need*, *prefer*, *want*, *wish*. These verbs differ as to the morphological realization of *for* in their infinitival complements:

(i) a. I {hate, intend, like, mean, prefer, wish} for/ϕ you to do that.
b. I {expect, need, want} *for/ϕ you to do that.

While *hate*, *intend*, *like mean*, *prefer*, and *wish* allow *for* to appear in their infinitival complements in addition to ϕ, *expect*, *need*, and *want* allows only ϕ. In our terms, this means that the verbs in (ia) can select full-fledged *c*P-CP, whereas the verbs in (ib) can only select *c*P without CP as their idiosyncratic properties when the subordinate interpretation can be provided through covert *c*-incorporation. However, if an adverbial element intervenes between the matrix verb and the complement, *for* must appear:

(ii) I want very much for/* ϕ you to do that.

This phenomenon is strongly parallel to the obligatoriness of *that* in sentences like *I believed wholeheartedly that/* ϕ you were right*. It may well be that the

intervening adverbials block a PF adjacency condition, though we will not go into the specific mechanism here.

[7] Apparent counterexamples such as (i) are found in the diachronic part of the Helsinki Corpus:

(i) ...nyston hwet heom to donne wære.

In this sentence, however, *whet heom to donne* is the complement of *wære* rather than *nyston*; thus it may well correspond to the construction *they didn't know what he was to do* in present-day English and may be analyzed as follows:

(ii) nyston [$_{cP}$ hwet$_i$ [$_{IP}$ heom$_j$ [$_{VP}$ t_j [$_{cP}$ to donne t_i] wære]]]

Seven examples of *wh + to* infinitives in OE I have found in the Helsinki Corpus are all followed by *be*. Therefore, the argument in the text is not affected.

[8] As an anonymous reviewer of *Studies in Modern English* points out, the structure in (36) has an apparent difficulty in explaining the [Obj to V-*enne*] order in OE. Then I stipulate that *to* optionally cliticizes to the left of V across objects and other elements in infinitival clauses. See Henry (1995) for a relevant discussion on cliticization of the infinitival complementizer.

[9] Among Postal's (1974) W-verbs listed in note 6, *hate*, *like*, *mean*, *need* and *wish* appeared in OE. If simple *c*P lacking CP occurs in the complement of these verbs, it is expected that V-*enne* without *to* would have been permissible, because only *c* is necessary for Case-checking and the subordinate interpretation would be assigned through covert *c*-incorporation as discussed in section 5.2. This prediction, however, is borne out and *to* was obligatorily present before V-*enne* in the complement of these verbs. I speculate that this is most plausibly due to the prepositional nature of *to* in OE. It is generally true that prepositions cannot be deleted even if PP is a complement of a predicate. If the infinitive marker *to* retained its original prepositional meanings in OE, its undeletability is also expected, though I will not go into the specific mechanism here.

[10] Since *for* and *to* are independent words with their own morphological forms, Merger operation is optional in *for to* infinitives in early ME, whereas this

operation is obligatory if morphological forms correspond to the *c*+C complex, like *to* in (38a) and *that*, *whether*, and *for* in present-day English.

[11] However, the following *wh* + *for to* infinitive constitutes an apparent counterexample to the structure of *for to* infinitives in (42a):

(i) The kyng of Englande ... wyst nat *where for to passe* ...
the king of England knew not where for to passe

(*Froiss* I CXXVi.152, *OED*)

If *for* realizes *c*+C, (i) is predicted to be excluded by the doubly-filled Comp filter in the same way as *wh* + *to* infinitives in OE. In the face of this fact, I suggest that the following structure was also available:

(ii) $[_{cP}$ Wh *c*(+wh) $[_{CP}$ C $[_{IP}$ for to $[_{VP}$ V-en]]]]

This indicates that there was an option in which *for to* complex was reanalyzed into Infl as a whole when *for* did not express directional or purposive meaning. As expected, *for to* infinitives could cooccur with raising verbs, whose complement is generally assumed to be IP:

(iii) And wel a lord he_i semede $[_{IP}$ t_i for to be]
and well a lord he seemed for to be

(Chaucer, *The Legend of Good Women*, Fischer (1988:71))

Watanabe (1993b) and Bošković (1997) assume that *for to* complex is base-generated in Infl and then *for* alone raises to C even in present-day English. However, I object to this view and suspect that (ii) is only a bridge phenomenon that died out in Modern English.

[12] After *let*, passive infinitives could appear as bare forms, as we have seen in (26) in the text. Let us suppose that a null counterpart of *to* heads the IP complement of *let*. As is well known, bare infinitival complements of causative *make* in present-day English can express indirect causation as well as direct causation:

(i) $[_{IP}$ John Infl made $[_{IP}$ Bill Infl do the work]]
↑<E> ↑<E>
(indirect causation)

I speculate that this is because a phonetically null Infl developed in the complement of *make* in a later period, perhaps in Modern English, by analogy with the complement of *let* headed by a null Infl.

[13] Since PRO is available at the level of UG, there is no acquisitional burden on the part of children. Tanaka (1994:92) argues that the infinitival affix functions as an argument if and only if it is Case-marked. If this is correct, then it might be that PRO in *to* infinitives first appeared in late ME, where dative Case of the infinitival affix was absent in cases like (42b), and later became obligatory when the affix was completely lost in the mid-16th century.

[14] Masayuki Ohkado (personal communication) points out that the example in (54a) may not necessarily imply that B-type verbs s-select propositions, since *probably* cannot occur in (i):

(i) *At first I did not see that she was probably so stingy.

However, I speculate that (i) is ungrammatical because the negative marker *not* takes scope over *probably*, which has an affirmative entailment. If so, the argument in the text can be maintained.

[15] Hirozo Nakano (personal communication) suggests that *that*- and ACI complements of B-type verbs differ in their referential abilities. While *that*-clause complements refer to intended facts in a possible world, ACI complements do not have a function to refer to particular facts. In spite of this difference, however, I use the term "proposition" in a broad sense and intend it to cover both of these cases.

[16] It is not clear how to derive the impossibility of A-movement across *c*P under the current minimalist model. Bošković (1997) argues that it is a consequence of Manzini's (1994) locality constraint on movement, which requires every movement proceed through the domain of each head. If the embedded subject raises to the matrix Case-checking position via intermediate Spec of *c*P, this movement will yield an improper chain, in which an A-bar-chain is followed by an A-chain. However, this approach cannot be maintained under the

definition of Attract F in Chomsky (1995) or our revised version, since the formal features of the embedded subject can directly move to the matrix verb skipping Spec of *c*P. It might be that an argument *c*P has a nominal feature and this feature blocks the movement of the nominal [D] feature of the embedded subject, which is included in the set of formal features carried along in covert raising for Case-checking. If this is correct, the impossibility of A-movement across *c*P will be accounted for in terms of the MLC in the sense of Chomsky (1995).

[17] Van Gelderen (1993) also suggests the relevance of Infl within infinitives to the ACI construction. The difference between our analysis and hers is that she claims that the introduction of Infl made all ACI constructions permissible in late ME including causative verb complements. Van Gelderen argues that causative verbs in OE had *persuade*-type three-place argument structures. It is unlikely, however, that the postverbal NP with causative verbs was their direct object rather than the subject of infinitives. For example, her analysis cannot explain the fact that the postverbal NP could be inanimate, which was not the case with *persuade*-type verbs:

(i) And treowa he deð færlice blowan
And trees he causes suddenly bloom-Inf
'and he causes the trees to bloom suddenly'
(Napier *HomU*. 34 109, Fischer (1989:189))

The explanation in the text has an advantage of being able to analyze the infinitival complement in (i) as bare VP.

[18] In OE, control *to* infinitives that appeared as the complement of B-type verbs could represent unrealized events (see (18c) in the text). Note that since there is no Infl node in OE infinitives, the [±tense] feature is irrelevant. The future orientation of OE *to* infinitives is most plausibly due to the directional and purposive meanings of the complementizer *to*.

[19] It must be postulated that in the FP construction where there is no argument DP that c-commands the embedded predicate, an infinitival affix that functions as an external argument can saturate a syntactic predicate.

Chapter 6

Conclusion and Further Issues

6.1. Grand Summary

In this thesis I have discussed the mechanism of parametric changes in the history of English. I started off from the functional parametrization hypothesis, which states that parameters of the language faculty is confined to the domain of functional categories in the lexicon, and asked if this hypothesis is valid for an explanation of diachronic changes as well. Specifically, I raised the following questions in Chapter 1: (i) what functional categories should be posited in the phrase structure; and (ii) whether there are any generalizations or principles of parameter resetting, and if any, how they should be formulated.

In Chapter 2, I considered question (i) with special reference to the structure of complementizers as a groundwork for later explanations of diachronic changes. I proposed the split-CP hypothesis, in which complementizers pronounced as single words such as *that* and *whether* are lexically decomposed into two distinct

categories, *c* and C. I demonstrated that this hypothesis successfully accounts for *that*-deletion and the *that*-trace effect in present-day English. It was also argued that the lower C is a pure subordinator and it is not included in main clauses; hence, the root/non-root asymmetry with respect to the basic clause structure was obtained.

In Chapter 3, I tackled question (ii) by analyzing the development of the complementizer system in English. The answer was that there are no principles of language change but still a generalization holds. I claimed that parameter resetting can be described in terms of diachronic feature shift among functional categories and that the process of feature shift exhibits the property of locality. This property enabled us to account for the simultaneous demise of doubly-filled Comp and V2 in the mid-15th century and the loss of V-to-I movement and the emergence of the *that*-trace effect in the 16th century.

In Chapter 4, I attempted to derive the effect of the locality property from principles of UG. The definition of Attract F was revised so as to incorporate the notion of Attraction Domain, which prescribes the correlations between the richness of verbal morphology and the domains where affixes are visible to the attractor. Given this more articulated theory of head movement, the loss of verb movement is ultimately attributed to the decline of verbal inflections, which is essentially a matter of language use. Thus, we can dispense with independent principles of language change to explain the parameter resetting concerning verb movement.

In Chapter 5, I examined the development of infinitival constructions to show that diachronic feature shift is observed in non-finite clauses as well and thereby demonstrate that the proposed mechanism has a wide range of explanatory power. As in finite clauses, parametric changes within infinitives were induced by the decline of the infinitival affix together with functional changes of complementizers. On the other hand, a characteristic of the feature shift within infinitives was that it was an upward shift from a verbal head to Infl and the Infl node was newly introduced into the phrase structure of infinitives as a result of the shift.

Throughout this thesis, I regarded parameter resetting in both finite and infinitival clauses as positional changes of certain features (i.e. [+F] and [−finite]). This view contrasts with a standard assumption in the minimalist model that parametric differences are reduced to strength of features. Chomsky (1995) maintains that the verbal feature in Infl is strong in languages with overt V-to-I movement whereas it is weak in languages without this operation. However, it is difficult to account for language change in terms of the dichotomy of strong/weak features, since if language acquisition is properly carried out, there will be no chance that the value of a feature changes from the strong setting to the weak one. Under our feature shift hypothesis, by contrast, this kind of logical problem does not arise, because diachronic feature shift can take place without changing the surface word order of PLD. Thus, it is not the strength of features but their positions that change in parameter resetting.

6.2. Further Issues

There are a number of parametric changes that I could not deal with in this thesis. For example, I paid little attention to the change of the basic word order from OV to VO that occurred around 1200. Kayne (1994) proposes a restrictive phrase structure theory which states that directionality plays no role in the computational system. In this antisymmetric view of syntax, all languages are universally head-initial and the head parameter does not exist. The OV order is derived from the basic VO order as a result of overt object shift. Roberts (1997) and van der Wurff (1997) put forth analyses of this change on the basis of Kayne's theory. The structures of the OV and VO word orders they propose are represented as follows:

(1) a. $[_{\text{AgrOP}}$ Obj V+AgrO $[_{\text{VP}}\ t_{\text{subj}}\ t_{\text{V}}\ t_{\text{obj}}\ ]]$ (OV)

b. $[_{\text{AgrOP}}$ AgrO $[_{\text{VP}}\ t_{\text{subj}}$ V Obj $]]$ (VO)

Roberts and van der Wurff both presuppose the projection of AgrO. In (1a), the object overtly raises to [Spec, AgrO], and V also adjoins to the head AgrO to check off the Case feature of the object. As a result, the surface OV order is obtained. On the other hand, the VO order reflects the base positions of V and the object, as shown in (1b). Thus, under Kayne's system, the change from OV to VO entails the loss of V-to-AgrO movement.

It is not clear, however, whether this analysis can be readily paraphrased in terms of the framework advanced in this thesis. It was argued that the shift of [+F] from *c* to Infl in the latter half of the

15th century caused the demise of V2 and the subsequent shift from Infl to *v* in early Modern English led to the loss of V-to-I movement. The hypothesis that V ceased to raise to AgrO around 1200 seems to go against unidirectionality of the loss of verb movement. It might be that the feature that motivates V-to-AgrO movement is not [+F] but some other independent feature of AgrO. Alternatively, a more promising approach is to hypothesize that there is no verb movement involved at all in overt object shift. Since we have adopted the basic phrase structure which dispenses with Agr, overt object shift will be simply described as movement of an object into [Spec, *v*P] without accompanying verb raising. Though it seems worth pursuing analyses of the word order change and other syntactic developments in the history of English within the present theoretical framework, I leave to a future research program to explore these areas.

References

Aarts, Bas (1992) *Small Clauses in English: The Nonverbal Types*, Mouton de Gruyter, Berlin.

Allen, Cynthia (1977) *Topics in Diachronic English Syntax*, Doctoral dissertation, University of Massachusetts, Amherst. [Published by Garland, New York, 1980.]

Amano, Masachiyo (1994) "Henkei Seisei Bunpo Riron ni okeru Jodoshi *Do* no Kenkyu-shi (Studies of Auxiliary *Do* in the History of the Transformational Generative Grammar)," in Nakao, Yuji and Masachiyo Amano, eds. (1994), 54-78.

Aoun, Joseph (1985) *A Grammar of Anaphora*, MIT Press, Cambridge, MA.

Araki, Kazuo and Masatomo Ukaji (1984) *Eigoshi IIIA* (History of English IIIA), Taishukan, Tokyo.

Arimoto, Masatake (1994) "Bun no Kozo to Meirei Bun no *Do* (Sentence Structures and *Do* in Imperative Sentences)," in Nakao, Yuji and Masachiyo Amano, eds. (1994), 247-280.

Arimura, Kaneaki (1998) "Sosei Shogo to Setuji: Jodoshi no Sintakusu (Feature Checking and Affixes: The Syntax of Auxiliaries)," *Eigo Seinen* 144.4, 2-6.

Arnold, Mark (1995) *Case, Periphrastic* Do *and the Loss of Verb Movement in English*, Doctoral dissertation, University of

Maryland.

Authier, J.-Marc (1992) "Iterated CPs and Embedded Topicalization," *Linguistic Inquiry* 23, 329-336.

Bach, Emmon (1977) "Review Article: *On Raising: One Rule of English Grammar and Its Theoretical Implications*, by Paul Postal, MIT Press, Cambridge, MA, 1974," *Language* 53, 621-654.

Baker, Mark (1988) *Incorporation: A Theory of Grammatical Function Changing*, University of Chicago Press, Chicago.

Baker, Mark, Kyle Johnson, and Ian Roberts (1989) "Passive Arguments Raised," *Linguistic Inquiry* 20, 219-251.

Bauer, Brigitte (1995) *The Emergence and Development of SVO Patterning in Latin and French: Diachronic and Psycholinguistic Perspectives*, Oxford University Press, New York.

Bayer, Josef (1984) "Comp in Bavarian Syntax," *The Linguistic Review* 3, 209-274.

Beal, Joan (1988) "Goodbye to All *That*? The History and Present Behaviour of Optional *That*," in Nixon, Graham and John Honey, eds. (1988), 49-66.

Belletti, Adriana and Luigi Rizzi, eds. (1996) *Parameters and Functional Heads: Essays in Comparative Syntax*, Oxford University Press, New York.

Bergh, Gunnar and Aimo Seppänen (1994) "Subject Extraction in English: The Use of the *That*-Complementizer," *English*

Maryland.

Authier, J.-Marc (1992) "Iterated CPs and Embedded Topicalization," *Linguistic Inquiry* 23, 329-336.

Bach, Emmon (1977) "Review Article: *On Raising: One Rule of English Grammar and Its Theoretical Implications*, by Paul Postal, MIT Press, Cambridge, MA, 1974," *Language* 53, 621-654.

Baker, Mark (1988) *Incorporation: A Theory of Grammatical Function Changing*, University of Chicago Press, Chicago.

Baker, Mark, Kyle Johnson, and Ian Roberts (1989) "Passive Arguments Raised," *Linguistic Inquiry* 20, 219-251.

Bauer, Brigitte (1995) *The Emergence and Development of SVO Patterning in Latin and French: Diachronic and Psycholinguistic Perspectives*, Oxford University Press, New York.

Bayer, Josef (1984) "Comp in Bavarian Syntax," *The Linguistic Review* 3, 209-274.

Beal, Joan (1988) "Goodbye to All *That*? The History and Present Behaviour of Optional *That*," in Nixon, Graham and John Honey, eds. (1988), 49-66.

Belletti, Adriana and Luigi Rizzi, eds. (1996) *Parameters and Functional Heads: Essays in Comparative Syntax*, Oxford University Press, New York.

Bergh, Gunnar and Aimo Seppänen (1994) "Subject Extraction in English: The Use of the *That*-Complementizer," *English*

Historical Linguistics 1992, ed. by Francisco Fernández, Miguel Fuster, and Juan Calvo, 131-143, John Benjamins, Amsterdam.

Berwick, Robert (1985) *The Acquisition of Syntactic Knowledge*, MIT Press, Cambridge, MA.

Besten, Hans den (1977) "On the Interaction of Root Transformations and Lexical Deletive Rules," *On the Formal Syntax of Westgermania*, ed. by W. Abraham, 47-131, John Benjamins, Amsterdam.

Besten, Hans den and Corretje Moed-van Walraven (1986) "The Syntax of Verbs in Yiddish," in Haider, Hubert and Martin Prinzhorn, eds. (1986), 111-135.

Bhatt, Rakesh and James Yoon (1991) "On the Composition of Comp and Parameters of V2," *WCCFL* 10, 41-52.

Bobaljik, Jonathan (1995) *Morphosyntax: The Syntax of Verbal Inflection*, Doctoral dissertation, MIT.

Boser, K., B. Lust, L. Santelmann, and J. Whitman (1992) "The Syntax of CP and V-2 in Early Child German (ECG): The Strong Continuity Hypothesis," *NELS* 22, 51-65.

Bošković, Željko (1997) *The Syntax of Nonfinite Complementation: An Economy Approach*, MIT Press, Cambridge, MA.

Bresnan, Joan (1972) *Theory of Complementation in English Syntax*, Doctoral dissertation, MIT. [Published by Garland, New York, 1979.]

Campos, Héctor and Paula Kempchinsky, eds. (1995) *Evolution and Revolution in Linguistic Theory*, Georgetown University Press,

Washington, D.C.

Carroll, Susanne (1983) "Remarks on *For-To* Infinitives," *Linguistic Analysis* 12, 415-451.

Chomsky, Noam (1977) *Essays on Form and Interpretation*, North-Holland, New York.

Chomsky, Noam (1981) *Lectures on Government and Binding*, Foris, Dordrecht.

Chomsky, Noam (1986a) *Knowledge of Language: Its Nature, Origin, and Use*, Praeger, New York.

Chomsky, Noam (1986b) *Barriers*, MIT Press, Cambridge, MA.

Chomsky Noam (1991) "Some Notes on Economy of Derivation and Representation," in Freidin, Robert, ed. (1991), 417-454.

Chomsky, Noam (1993) "A Minimalist Program for Linguistic Theory," in Hale, Kenneth and Jay Keyser, eds. (1993), 2-52.

Chomsky, Noam (1994) *Bare Phrase Structure, MIT Occasional Papers in Linguistics* 5. [Reprinted in Campos, Héctor and Paula Kempchinsky, eds. (1995), 51-109.]

Chomsky, Noam (1995) "Categories and Transformations," *The Minimalist Program*, 219-394, MIT Press, Cambridge, MA.

Chomsky, Noam (1998) "Minimalist Inquiries: The Framework," ms., MIT.

Chomsky, Noam and Howard Lasnik (1977) "Filters and Control," *Linguistic Inquiry* 8, 425-504.

Clark, Robin and Ian Roberts (1993) "A Computational Model of Language Learnability and Language Change," *Linguistic*

Inquiry 24, 299-345.

Collins, Chris (1997) *Local Economy*, MIT Press, Cambridge, MA.

Culicover, Peter (1991) "Topicalization, Inversion, and Complementizers in English," ms., The Ohio State University.

Curme, George (1931) *Syntax*, D. C. Heath, Boston.

Denison, David (1993) *English Historical Syntax*, Longman, London.

Doherty, Cathal (1997) "Clauses without Complementizers: Finite IP-Complementation in English," *The Linguistic Review* 14, 197-220.

Dubinsky, Stanley and Kemp Williams (1995) "Recategorization of Prepositions as Complementizers: The Case of Temporal Prepositions in English," *Linguistic Inquiry* 26, 125-137.

Ellegård, Alvar (1953) *The Auxiliary* Do*: The Establishment and Regulation of Its Use in English*, Almqvist and Wiksell, Stockholm.

Emonds, Joseph (1976) *A Transformational Approach to English Syntax: Root, Structure-Preserving, and Local Transformations*, Academic Press, New York.

Fischer, Olga (1988) "The Rise of the *For* NP *to* V Construction: An Explanation," in Nixon, Graham and John Honey, eds. (1988), 67-88.

Fischer, Olga (1989) "The Origin and the Spread of the Accusative and Infinitive Construction in English," *Folia Linguistica Historica* 8, 143-217.

Fischer, Olga (1992) "Syntactic Change and Borrowing: The Case of

the Accusative-and-Infinitive Construction in English," *Internal and External Factors in Syntactic Change*, ed. by Marinel Gerritsen and Diete Stein, 17-88, Mouton de Gruyter, Berlin.

Fischer, Olga (1995) "The Distinction between *To* and Bare Infinitival Complements in Late Middle English," *Diachronica* 12, 1-30.

Freidin, Robert, ed. (1991) *Principles and Parameters in Comparative Grammar*, MIT Press, Cambridge, MA.

Freidin, Robert and Rex Sprouse (1991) "Lexical Case Phenomena," in Freidin, Robert, ed. (1991), 392-416.

Fukui, Naoki (1988) "Deriving the Difference between English and Japanese: A Case Study in Parametric Syntax," *English Linguistics* 5, 249-270.

Fukui, Naoki (1995) "The Principles-and-Parameters Approach: A Comparative Syntax of English and Japanese," *Approaches to Language Typology*, ed. by Masayoshi Shibatani and Theodora Bynon, 327-372, Clarendon Press, Oxford.

Geis, Michael (1970) *Adverbial Subordinate Clauses in English*, Doctoral dissertation, MIT.

Gelderen, Elly van (1993) *The Rise of Functional Categories*, John Benjamins, Amsterdam.

Gerritsen, Marinel (1984) "Divergent Word Order Developments in Germanic Languages: A Description and a Tentative Explanation," *Historical Syntax*, ed. by Jacek Fisiak, 107-135, Mouton, Berlin.

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

Grimshaw, Jane (1990) *Argument Structure*, MIT Press, Cambridge, MA.

Grimshaw, Jane (1991) "Extended Projection," ms., Brandeis University.

Grimshaw, Jane (1994) "Minimal Projection, Heads, and Optimality," ms., Rutgers University.

Grimshaw, Jane and Armin Mester (1988) "Light Verbs and θ-Marking," *Linguistic Inquiry* 19, 205-232.

Haegeman, Liliane (1991) *Introduction to Government and Binding Theory*, Blackwell, Oxford.

Haegeman, Liliane (1992) *Theory and Description in Generative Syntax: A Case Study in West Flemish*, Cambridge University Press, Cambridge.

Haider, Hubert and Martin Prinzhorn, eds. (1986) *Verb Second Phenomena in Germanic Languages*, Foris, Dordrecht.

Hale, Ken (1983) "Warlpiri and the Grammar of Non-configurational Languages," *Natural Language and Linguistic Theory* 1, 5-47.

Hale, Ken and Samuel Jay Keyser (1993) "On Argument Structure and the Lexical Expression of Syntactic Relations," in Hale, Ken and Samuel Jay Keyser, eds. (1993), 53-109.

Hale, Ken and Samuel Jay Keyser, eds. (1993) *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, MIT Press, Cambridge, MA.

Halle, Morris and Alec Marantz (1993) "Distributed Morphology and the Pieces of Inflection," in Hale, Ken and Samuel Jay Keyser, eds. (1993), 111-176.

Hamamatsu, Junji (1996) "On Doubly-Filled CP and DP Constructions in the History of English," *Studies in Modern English* 12, 1-16.

Hellan, Lars and Kirsti Christensen, eds. (1986) *Topics in Scandinavian Syntax,* D. Reidel, Dordrecht.

Henry, Alison (1995) *Belfast English and Standard English: Dialect Variation and Parameter Setting*, Oxford University Press, New York.

Higginbotham, James (1985) "On Semantics," *Linguistic Inquiry* 16, 547-593.

Hoekstra, Eric (1993) "On the Parametrisation of Functional Projections in CP," *NELS* 23, 191-204.

Holmberg, Anders and Christer Platzack (1995) *The Role of Inflection in Scandinavian Syntax*, Oxford University Press, Oxford.

Hornstein, Norbert (1990) *As Time Goes By: Tense and Universal Grammar*, MIT Press, Cambridge, MA.

Hornstein, Norbert and David Lightfoot (1987) "Predication and PRO," *Language* 63, 23-52.

Hornstein, Norbert and David Lightfoot (1991) "On the Nature of Lexical Government," in Freiden, Robert, ed. (1991), 365-391.

Huang, C.-T. James (1982) *Logical Relations in Chinese and the Theory of Grammar*, Doctoral dissertation, MIT.

Hulk, Aafke and Ans van Kemenade (1997) "Negation as a Reflex of Clause Structure," *Negation and Polarity: Syntax and Semantics*, ed. by Danielle Forget, Paul Hirschbühler, and France Martineau, John Benjamins, Amsterdam.

Ishikawa, Kazuhisa (1995) "A History of *Not*: The Change from a Phrase to a Head," *English Linguistics* 12, 197-221.

Itô, Junko and Armin Mester (1995) "The Core-Periphery Structure of the Lexicon and Constraints on Reranking," *University of Massachusetts Occasional Papers* 18, 181-209.

Jespersen, Otto (1909-1949) *A Modern English Grammar on Historical Principles*, 7 vols, Munksgaard, Copenhagen, Carl Winter, Heidelgerg, George Allen and Unwin, London.

Jespersen, Otto (1967) *Growth and Structure of the English Language*, 9th ed., Basil Blackwell, Oxford.

Johnson, Kyle (1988) "Clausal Gerunds, the ECP, and Government," *Linguistic Inquiry* 19, 583-609.

Jonas, Dianne (1996) "Clause Structure, Expletives and Verb Movement," *Minimal Ideas: Syntactic Studies in the Minimalist Framework*, ed. by Werner Abraham, Samuel Epstein, Höskuldur Thráinsson, and Jan-Wouter Zwart, 167-188, John Benjamins, Amsterdam.

Kageyama, Taro (1992) "Agr in Old English *To*-Infinitives," *Lingua* 88, 91-128.

Kato, Kozo (1997) "Clause-Internal *Nu* in Old English: A Root/Non-root Asymmetry," *Linguistics and Philology* 17, 15-39.

Kayne, Richard (1975) *French Syntax: The Transformational Cycle*, MIT Press, Cambridge, MA.

Kayne, Richard (1984) *Connectedness and Binary Branching*, Foris, Dordrecht.

Kayne, Richard (1994) *The Antisymmetry of Syntax*, MIT Press, Cambridge, MA.

Kemenade, Ans van (1987) *Syntactic Case and Morphological Case in the History of English*, Foris, Dordrecht.

Kemenade, Ans van (1997) "V2 and Embedded Topicalization in Old and Middle English," in Kemenade, Ans van and Nigel Vincent, eds. (1997), 326-352.

Kemenade, Ans van and Nigel Vincent, eds. (1997) *Parameters of Morphosyntactic Change*, Cambridge University Press, Cambridge.

Kiparsky, Paul and Carol Kiparsky (1970) "Fact," *Progress in Linguistics*, ed. by Manfred Bierwisch and Karl Heidolph, 143-173, Mouton, The Hague.

Kroch, Anthony (1989) "Reflexes of Grammar in Patterns of Language Change," *Language Variation and Change* 1, 199-244.

Kroch, Anthony and Ann Taylor (1997) "Verb Movement in Old and Middle English: Dialect Variation and Language Contact," in Kemenade, Ans van and Nigel Vincent, eds. (1997), 297-325.

Larson, Richard (1987) "Extraction and Multiple Selection in PP," *MIT Working Papers in Linguistics* 9, 119-135.

Lasnik, Howard (1981) "Restricting the Theory of Transformations: A Case Study," *Explanation in Linguistics: The Logical Problem of Language Acquisition*, ed. by Norbert Hornstein and David Lightfoot, 152-173, Longman, London.

Lasnik, Howard (1992) "Case and Expletives: Notes toward a Parametric Account," *Linguistic Inquiry* 23, 381-405.

Lasnik, Howard (1995) "Verbal Morphology: *Syntactic Structures* Meets the Minimalist Program," in Campos, Héctor and Paula Kempchinsky, eds. (1995), 251-275.

Lasnik, Howard and Mamoru Saito (1991) "On the Subject of Infinitives," *CLS* 27, Part 1: *The General Session*, 324-343.

Lasnik, Howard and Mamoru Saito (1992) *Move α: Conditions on Its Application and Output*, MIT Press, Cambridge, MA.

Lebeaux, David (1986) "The Interpretation of Derived Nominals," *CLS* 22, Part1: *Papers from the General Session*, 231-247.

Lightfoot, David (1979) *Principles of Diachronic Syntax*, Cambridge University Press, Cambridge.

Lightfoot, David (1991) *How to Set Parameters: Arguments from Language Change*, MIT Press, Cambridge, MA.

Lightfoot, David (1999) *The Development of Language: Acquisition, Change, and Evolution*, Blackwell, Oxford.

Manzini, Rita (1994) "Locality, Minimalism, and Parasitic Gaps," *Linguistic Inquiry* 25, 481-508.

Manzini, Rita and Kenneth Wexler (1987) "Parameters, Binding Theory, and Learnability," *Linguistic Inquiry* 18, 413-444.

Martin, Roger (1992) "On the Distribution and Case Features of PRO," ms., University of Connecticut.

McDavid, V. (1964) "The Alternation of *That* and Zero in Noun Clauses," *American Speech* 39, 102-113.

Melvold, Janis (1991) "Factivity and Definiteness," *MIT Working Papers in Linguistics* 15, 97-117.

Mitchell, Bruce (1985) *Old English Syntax*, 2 vols., Clarendon Press, Oxford.

Müller, Gereon and Wolfgang Sternefeld (1993) "Improper Movement and Unambiguous Binding," *Linguistic Inquiry* 24, 461-507.

Mustanoja, Tauno (1960) *A Middle English Syntax: Part I*, Société Néophilologique, Helsinki.

Myers, Scott (1984) "Zero-Derivation and Inflection," *MIT Working Papers in Linguistics* 7, 53-69.

Nakajima, Heizo (1996) "Complementizer Selection," *The Linguistic Review* 13, 143-164.

Nakano, Hirozo (1985) "Eigo Ho Jodoshi no Saibunseki ni tuite no Ichi Kosatu (An Analysis of Reanalysis of English Modal Auxiliaries)," *Modern English Studies* 2, 48-75.

Nakano, Hirozo (1994) "Jodoshi *Do* no Hattatu: Meidai-nai-teki Hyogen kara Meidai-teki Hyogen e (The Development of Auxiliary *Do*: From Proposition-Internal Expression to Propositional Expression)," in Nakao, Yuji and Masachiyo Amano, eds. (1994), 302-322.

Nakano, Hirozo (1996) "Teikei Setu no Imi Bunseki (A Semantic

Analysis of Finite Clauses)," *Nagoya Daigaku Bungakubu Kenkyu Ronsyu* 124, Bungaku 42, 79-88.

Nakano, Hirozo (1999) "Sense and Reference Grammar," class lecture material, Nagoya University.

Nakao, Toshio (1972) *Eigoshi II* (History of English II), Taishukan, Tokyo.

Nakao, Yuji and Masachiyo Amano, eds. (1994) *Jodoshi* Do*: Kigen, Hattatu, Kinou* (Auxiliary *Do*: Its Origin, Development, and Function), Eichosha, Tokyo.

Nawata, Hiroyuki (1998a) "A Φ-Feature Checking Approach to PRO," *Central Japan English Studies* 17, 83-97.

Nawata, Hiroyuki (1998b) "On the Development of Infinitival Constructions in the History of English," *Studies in Modern English* 14, 29-51.

Nawata, Hiroyuki (1998c) "The Distribution of Complementizers and the Split-CP Hypothesis Revisited," *Linguistics and Philology* 18, 73-94.

Nawata, Hiroyuki (1999a) "Locality of Diachronic Feature Shift and *Do*-Support," *Papers from the 16th National Conference of The English Linguistic Society of Japan*, 171-180.

Nawata, Hiroyuki (1999b) "Split-CP, Doubly-Filled Comp, and Locality of Language Change," *English Linguistics* 16, 121-144.

Nawata, Hiroyuki (forthcoming) "An Analysis of the *That*-Trace Effect in the History of English," *Papers from the 17th National*

Conference of The English Linguistic Society of Japan.

Nixon, Graham and John Honey, eds. (1988) *An Historic Tongue: Studies in English Linguistics in Memory of Barbara Strang*, Routledge, London.

Ono, Shigeru, and Toshio Nakao (1980) *Eigoshi I* (History of English I), Taishukan, Tokyo.

Ormazabal, Javier (1994) "PRO, Null Case and the Interpretation of Complements," *NELS* 24, 475-489.

Otsu, Norihiko (1999) "Koki Tyu-eigo ni okeru Setuzokushi *That* no Umu ni tuite (On the Absence of the Conjunction *That* in Late Middle English)," *Papers from the 16th National Conference of The English Linguistic Society of Japan*, 191-200.

Otsu, Yukio (1989) "Shinri Gengo-gaku, (Psycholinguistics)" in *Eigo-Gaku no Kanren Bunya* (Fields Related to English Linguistics), by Masayoshi Shibatani, Yukio Otsu, and Aoi Tsuda, 181-361, Taishukan, Tokyo.

Pesetsky, David (1992) "Zero Syntax II: An Essay on Infinitives," ms., MIT.

Pesetsky, David (1995) *Zero Syntax: Experiencers and Cascades*, MIT Press, Cambridge, MA.

Platzack, Christer (1986) "Comp, Infl, and Germanic Word Order," in Hellan, Lars and Kirsti Christensen, eds. (1986), 185-234.

Pollock, Jean-Yves (1989) "Verb Movement, Universal Grammar, and the Structure of IP," *Linguistic Inquiry* 20, 365-424.

Postal, Paul (1974) *On Raising: One Rule of English Grammar and*

Its Theoretical Implications, MIT Press, Cambridge, MA.

Potsdam, Eric (1996) *Syntactic Issues in the English Imperatives*, Doctoral dissertation, University of California, Santa Cruz. [Published by Garland, New York, 1998.]

Potsdam, Eric (1997) "English Verbal Morphology and VP Ellipsis," *NELS* 27, 353-368.

Pyles, Thomas (1964) *The Origins and Development of the English Language*, Harcourt, Brace and World, New York.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (1985) *A Comprehensive Grammar of the English Language*, Longman, London.

Rappaport, Malka Hovav and Beth Levin (1992) "*-Er* Nominals: Implications for the Theory of Argument Structure," *Syntax and Semantics* 26: *Syntax and the Lexicon*, ed. by Tim Stowell and Eric Wehrli, 127-153, Academic Press, New York.

Rissanen, Matti (1991) "On the History of *That*/Zero as Object Clause Links in English," *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, ed. by Karin Aijmer and Bengt Altenberg, 272-289, Longman, London.

Rissanen, Matti (1997) "Optional *That* with Subordinators in Middle English," *Language History and Linguistic Modelling: A Festschrift for Jacek Fisiak on His 60th Birthday*, ed. by Raymond Hickey and Stanislaw Puppel, 373-383, Mouton de Gruyter, Berlin.

Rizzi, Luigi (1982) *Issues in Italian Syntax*, Foris, Dordrecht.

Rizzi, Luigi (1986) "Null Objects in Italian and the Theory of *pro*," *Linguistic Inquiry* 17, 501-557.

Rizzi, Luigi (1990) *Relativized Minimality*, MIT Press, Cambridge, MA.

Rizzi, Luigi (1991) "Residual Verb Second and the *Wh*-Criterion," *Technical Reports in Formal and Computational Linguistics* 2, University of Geneva. [Reprinted in Belletti, Adriana and Luigi Rizzi, eds. (1996), 63-90.]

Rizzi, Luigi (1996) "The Fine Structure of the Left Periphery," *Elements of Grammar: Handbook of Generative Syntax*, ed. by Liliane Haegeman, Kluwer, Dordrecht.

Rizzi, Luigi and Ian Roberts (1989) "Complex Inversion in French," *Probus* 1, 1-30. [Reprinted in Belletti, Adriana and Luigi Rizzi, eds. (1996), 91-116.]

Roberts, Ian (1993) *Verbs and Diachronic Syntax: A Comparative History of English and French*, Kluwer, Dordrecht.

Roberts, Ian (1997) "Directionality and Word Order Change in the History of English," in Kemenade, Ans van and Nigel Vincent, eds. (1997), 397-426.

Roberts, Ian (1998a) "Verb Movement and Markedness," *Language Creation and Change: Creolization, Diachrony and Development*, ed. by Michel DeGraff, MIT Press, Cambridge, MA.

Roberts, Ian (1998b) "*Have/Be* Raising, Move F, and Procrastinate," *Linguistic Inquiry* 29, 113-125.

Rohrbacher, Bernhard (1994) *The Germanic Languages and the Full Paradigm: A Theory of V to I Raising*, Doctoral dissertation, University of Massachusetts, Amherst.

Rothstein, Susan (1995) "Pleonastics and the Interpretation of Pronouns," *Linguistic Inquiry* 26, 499-529.

Safir, Ken (1993) "Perception, Selection, and Structural Economy," *Natural Language Semantics* 2, 47-70.

Shima, Etsuro (1999) "Two Types of *Wh*-features," *Lingua* 107, 189-206.

Stowell, Tim (1981) *Origins of Phrase Structure*, Doctoral dissertation, MIT.

Stowell, Tim (1982) "The Tense of Infinitives," *Linguistic Inquiry* 13, 561-570.

Strang, Barbara (1970) *A History of English*, Methuen, London.

Takano, Yuji (1996) *Movement and Parametric Variation in Syntax*, Doctoral dissertation, University of California, Irvine.

Tanaka, Tomoyuki (1994) "On the Realization of External Arguments in Infinitives," *English Linguistics* 11, 76-99.

Tanaka, Tomoyuki (1997) "Minimalism and Language Change: The Historical Development of To-Infinitives in English," *English Linguistics* 14, 320-341.

Tanaka, Tomoyuki and Harumasa Miyashita (1999) "A Note on *For To* Infinitives in Belfast English and Middle English," *English Linguistics* 16, 501-507.

Taraldsen, Knut (1986) "*Som* and the Binding Thoery," in Hellan,

Lars and Kirsti Christensen, eds. (1986), 149-184.

Terzi, A. (1992) *PRO in Finite Clauses: A Study of the Inflectional Heads of the Balkan Languages*, Doctoral dissertation, City University of New York.

Thráinsson, Höskuldur (1986) "V1, V2, V3 in Icelandic," in Haider, Hubert and Martin Prinzhorn, eds. (1986), 169-194.

Traugott, Elizabeth (1982) "From Propositional to Textual and Expressive Meanings: Some Semantic-Pragmatic Aspects of Grammaticalization," *Perspectives on Historical Linguistics*, ed. by Winfred Lehmann and Yakov Malkiel, John Benjamins, Amsterdam.

Traugott, Elizabeth (1986) "From Polysemy to Internal Semantic Reconstruction," *BLS* 12, 539-550.

Traugott, Elizabeth (1989) "On the Rise of Epistemic Meanings in English: An Example of Subjectification in Semantic Change," *Language* 65, 31-55.

Traugott, Elizabeth (1992) "Syntax," *The Cambridge History of the English Language*, vol. I, *The Beginnings to 1066*, ed. by Richard Hogg, 168-289, Cambridge University Press, Cambridge.

Travis, Lisa (1984) *Parameters and Effects of Word Order Variation*, Doctoral dissertation, MIT.

Ukaji, Masatomo (1992) "*I not say*: Bridge Phenomenon in Syntactic Change," *History of Englishes: New Methods and Interpretations in Historical Linguistics*, ed. by Matti Rissanen,

Ossi Ihalainen, Terttu Nevalainen, and Irma Taavitsainen, 453-462, Mouton de Gruyter, Berlin .

Ukaji, Masatomo (1998) "A Short History of *Whether*," paper presented at symposium, Setuzokushi o Meguru Shomondai (Issues of Conjunctions), 16th National Conference of The English Linguistic Society of Japan.

Ura, Hiroyuki (1996) *Multiple Feature-Checking: A Theory of Grammatical Function Splitting*, Doctoral dissertation, MIT.

Vikner, Sten (1995) *Verb Movement and Expletive Subjects in the Germanic Languages*, Oxford University Press, New York.

Vikner, Sten (1997) "V^0-to-I^0 Movement and Inflection for Person in All Tenses," *The New Comparative Syntax*, ed. by Liliane Haegeman, 189-213, Longman, London.

Visser, Frederik Theodoor (1963-1973) *An Historical Syntax of the English Language*, 4 vols., E. J. Brill, Leiden.

Wakayama, Masayuki (1996) "Bunpo Genri no Yusensei to Kyotyosei: Shokaku Tochi Kobun no Baai (Priority and Cooperation in Principles)," *Central Japan English Studies* 15, 97-112.

Warner, Anthony (1982) *Complementation in Middle English and the Methodology of Historical Syntax: A Study of the Wyclifite Sermons*, Croom Helm, London.

Warner, Anthony (1993) *English Auxiliaries: Structure and History*, Cambridge University Press, Cambridge.

Warner, Anthony (1997) "The Structure of Parametric Change, and V-movement in the History of English," in Kemenade, Ans van

and Nigel Vincent, eds. (1997), 380-393.

Watanabe, Akira (1993a) "Larsonian CP Recursion, Factive Complements, and Selection," *NELS* 23, 523-537.

Watanabe, Akira (1993b) *Agr-based Case Theory and Its Interaction with the A-Bar System*, Doctoral dissertation, MIT.

Watanabe, Akira (1994) "The Role of Triggers in the Extended Split Infl Hypothesis: Unlearnable Parameter Settings," *Studia Linguistica* 48.2, 156-178.

Webelhuth, Gert (1992) *Principles and Parameters of Syntactic Saturation*, Oxford University Press, New York.

Wurff, Wim van der (1997) "Deriving Object-Verb Order in Late Middle English," *Journal of Linguistics* 33, 485-509.

Yamane, Noriko (1997) "Rekisiteki On'in Henka to Saitekisei Riron: Shuwa Shousitu no Baai (Historical Phonological Change and Optimality Theory: A Case Study on the Demise of Schwa)," *On'in Kenkyu: Riron to Jissen* (Essays on Phonology: Theoretical and Descriptive Approaches) ed. by On'inron Kenkyukai, 167-172, Kaitakusha, Tokyo.

Yanagi, Tomohiro (1999) "Verb Movement and the Historical Development of Perfect Constructions in English," *English Linguistics* 16, 436-464.

Yoon, James (1996) "Ambiguity of Government and the Chain Condition," *Natural Language and Linguistic Theory* 14, 105-162.

Zubizarreta, Maria (1987) *Levels of Representation in the Lexicon and*

in the Syntax, Foris, Dordrecht.

Zubritskaya, Katya (1995) "Markedness and Sound Change in OT," *NELS* 25, vol. 2, 249-264.

Zwart, Jan-Wouter (1993) *Dutch Syntax: A Minimalist Approach*, Doctoral dissertation, University of Groningen.

Primary Sources and Dictionaries

v. d. *The Parker Chronicle*, MS. A, in Plummer, C., ed. (1892-1899) *Two of the Saxon Chronicles Parallel*, 2 vols., Clarendon Press, Oxford.

v. d. *The Prologue to the Bible*, in Forshall, J., and F. Madden, eds. (1850) *The Holy Bible by John Wycliffe and His Followers*, Oxford University Press, Oxford.

1422-1509 Gairdner, J., ed. (1965) *The Paston Letters*, 6 vols., AMS Press Inc., New York.

a 1425 Hamelius, P., ed. (1919) *Mandeville's Travels*, Early English Text Society, 153.

a1425 *The Cloud of Unknowing*, in Hodgson, P., ed. (1958) *The Cloud of Unknowing and the Book of Privy Counselling*, Early English Text Society, 218.

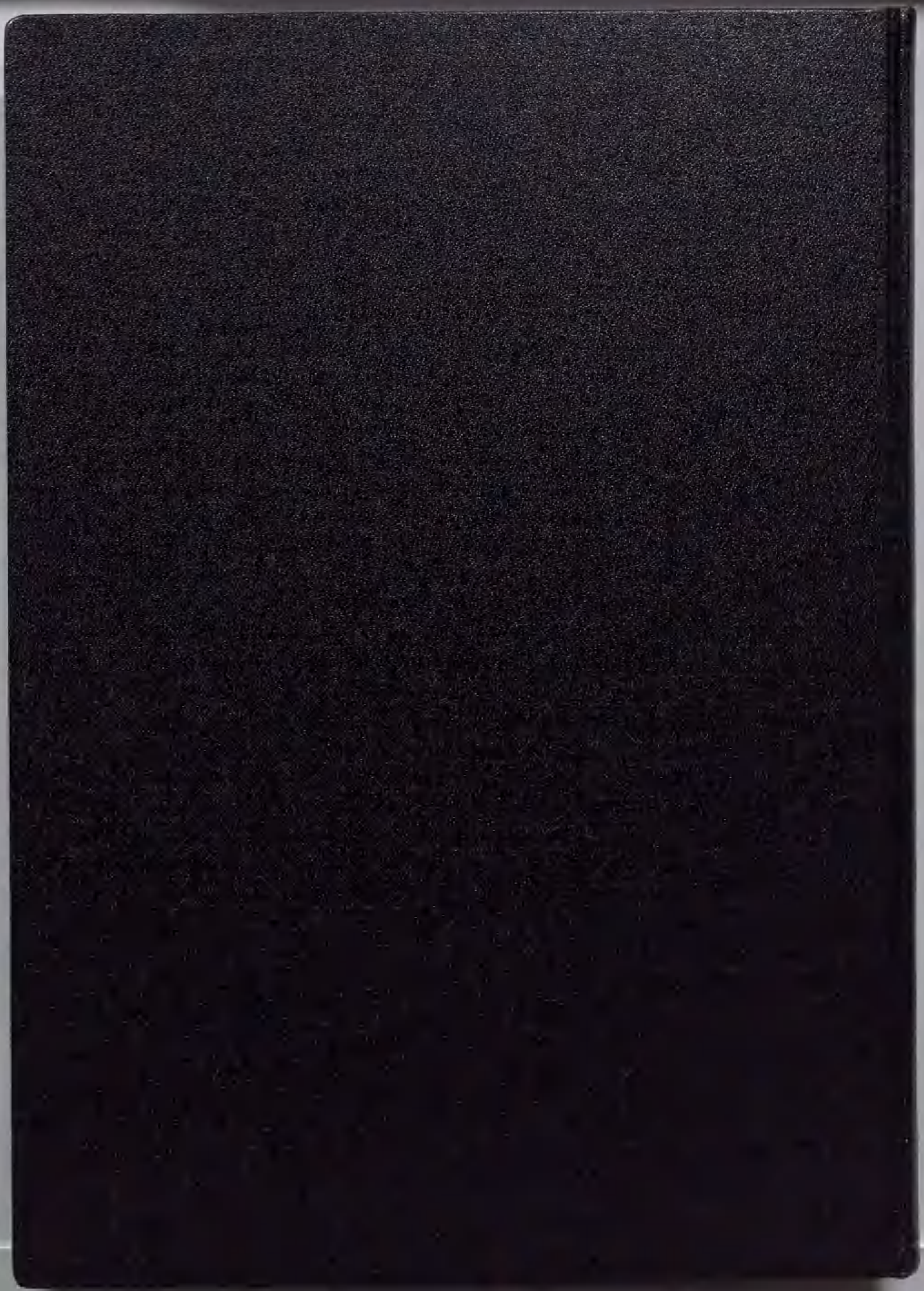
1607-1642 Gardiner, D., ed. (1933) *The Oxinden Letters*, Constable & Co. Ltd., London.

An Anglo-Saxon Dicitionary, based on the manuscript collections by J. Bosworth, edited and enlarged by T. N. Toller, 1898, Oxford University Press, London.

Middle English Dictionary (*MED*), ed. by H. Kurath and S. M. Kuhn, 1954-, University of Michigan Press, Michigan.

The Kenkyusha Dictionary of English Collocations, ed. by S. Ichikawa, et al., 1995, Kenkyusha, Tokyo.

The Oxford English Dictionary (*OED*), 2nd edition, prepared by J. A. Simpson and E. S. C. Weiner, 1989, Clarendon Press, Oxford.



inches 1 2 3 4 5 6 7 8

cm 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Kodak Color Control Patches

Blue Cyan Green Yellow Red Magenta White 3/Color Black

Kodak Gray Scale

C Y M

A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19