A HISTORY OF NOT: THE CHANGE FROM A PHRASE TO A HEAD

KAZUHISA ISHIKAWA

Aichi Gakuin University

This paper aims to study a syntactic change of not in the history of English. It is argued that not and its variants were generated in \([\text{Spec, NegP}]\) as XP categories in OE and ME, but they began to be reanalyzed as \(\text{Neg}^0\) in the sixteenth century. Noting that the OE and ME negative marker \(\text{ne}\) in \(\text{Neg}^0\) was missing in some cases, I assume \([\text{Neg e}]\) for \(\text{Neg}^0\) without \(\text{ne}\), and propose a licensing condition on \([\text{Neg e}]\). Then I claim that the condition, which was satisfied in OE and ME, began to be violated in the sixteenth century, and not came to fill in \([\text{Neg e}]\) to avoid the violation.*

1. Introduction

In the literature, various kinds of study have been made on the historical development of negative constructions in English. According to Jespersen's (1917, 1940) and Ukaji's (1984) observations, for instance, the typical development is schematically illustrated as follows:¹

* This is a revised version of the paper read at the 65th General Meeting of The English Literary Society of Japan held at Tokyo University on May 16, 1993. I am grateful to Hirozo Nakano, Masachiyo Amano, Katsuhide Sonoda, Hiroshi Terada, Mitsuru Maeda, Satomi Niwa, and two anonymous EL reviewers for their invaluable comments and suggestions on an earlier version of this paper.
¹ (1) only contains not and its variants which express sentential negation. In this paper, I will deal with this kind of negative marker, but not ones which express constituent negation as in (i).

(i) a. a not unapproachable figure
    b. Not always has she seasoned the meat.
Significantly, an EL reviewer points out that there are cases of sentential negation like (ii).

(ii) Not a single dress has she bought for ten years.
The combination Not a single can be viewed as an emphatic alternative to no as a countable determiner. Thus it is plausible to say that it is not Not but the combination Not a single (dress) that gives rise to the sentential negation: Not is only part of the sentential negative marker Not a single (dress). At this point, it seems to me that this kind of not needs another analysis. Here I will only deal with not and its variants which function as sentential negative markers by themselves.

English Linguistics 12 (1995) 197-221 — 197 —
© 1995 by the English Linguistic Society of Japan
Here it is noteworthy that *not* co-occurred with *do* in the sixteenth century, as in (1d). Some recent works attempt to account for *do*-support on the basis of the status of *not* as an X^0 category (cf. Chomsky (1991), Roberts (1993), etc.). Their common idea is that LF V^0 raising to Infl across the X^0 category *not* would give rise to a violation of the ECP and to avoid this, the insertion of *do* into Infl takes place instead of the V^0 raising. If this analysis of *not* is correct, it will be expected that *not* did not fall under an X^0 category until the sixteenth century.

In this paper, I will first show that *not* and its variants (henceforth simply *not* for expository convenience) fell under an XP category in OE and ME. Then, I will discuss the structure of OE and ME negative sentences under the NegP analysis. In the discussion, I will claim that *not* occupied [Spec, NegP] in OE and ME. Next I will show that *not* began to be generated under Neg^0 as an X^0 category in the sixteenth century. I will further examine why this change took place. Noting that the negative marker *ne* in Neg^0 was sometimes missing in the period from OE to LME, I will argue that [Neg e], Neg^0 with no lexical content, was allowed in this period. As a device of the grammar, I propose a licensing condition on [Neg e] as in the following:

(2) [Neg e] must be governed by strong Agr.

On the basis of (2) and the fact that Agr became weak in the sixteenth century (cf. Pollock (1989)), I will provide an account for the generation of *not* under Neg^0. I will further motivate (2) by examining facts from other languages. Finally, I will propose an identification principle for [Neg e] to recover its content from a phonetically realized negative marker such as *not*.

---

2 The parentheses indicate that negative markers such as *nawiht, naht* were not firmly established in OE.

3 Besides OE *nawiht/naht* and ME *not*, there were a lot of variants: OE *nowiht, noht, nauht, nawht*, etc.; ME *note, notte, nat, nate, no3t, nawt, nut, nought*, etc.
2. OE and ME *not* as an XP Category

In this section, it is shown that OE and ME *not* is an XP category. It is worth noting that OE and ME *not* is similar in distribution to the negative adverbs *never* and its variants (OE *næfre*; ME *nauwer, nefre*, etc.), which are XP categories. Let us begin by seeing the distribution of the former. In the typical patterns (1a) and (1b), *not* occurs immediately after a finite verb in clauses where the verb is not in the final position. The relevant examples are shown as in the following:

(3) a. Ac hi hi ne ladiað nowiht, ...
(Healey and Venezky (1980))

but they themselves NegM exculpate not
‘but they do not exculpate themselves, …’

b. It ne come no3t poru. (c 1325 Glo. Chron.A 4228)
It NegM came not through
‘It did not come through.’

(MED)

When a verb with the negative marker *ne* is preposed, OE and ME *not* usually appears immediately after a subject, if there are no pronomial objects.

(4) a. þa þa he waes farende to oðrum stowum, ne
when he was going to other places NegM
onwænde he noht þone ealdan feond,...
(Healey and Venezky (1980))

upset he not the old enemy
‘when he was on the way to other places, he did not defeat the old enemy, …’

b. Ne þearf þu noht dreden þe attri
NegM are required you not to dread the poisonous
neddre of helle. (c1230 Ancr. 71.20)

snake of hell
‘You need not dread the poisonous snake of hell.’

(MED)

OE *not* also occurs normally between a subject and the complement of

4 In (3) and the other relevant cases below, the gloss on *ne* is represented as NegM, which stands for a negative marker.
5 MED = *Middle English Dictionary*, University of Michigan Press, Ann Arbor, Michigan.
a verb in verb-final embedded clauses.

(5) ðonne wile he gereceeanðæt he noht unryhtlíc hit ne then will he explain that he not unrighteously it NegM forslæwde, ... (CP 39.283.24) hindered

‘then he will explain that he did not hinder it unrighteously, ...

(Healey and Venezky (1980))

Let us next turn to OE and ME never and its variants. They can occupy the same positions as not can, as illustrated below. The examples (6), (7), and (8) correspond to (3), (4), and (5), respectively.

(6) a. ac hi ne geendiað næfre. (ÆLS (Christmas) 30)

but they NegM end never

‘but they never come to an end.’ (Healey and Venezky (1980))

b. ðat dai hie ne com nauwer ute.

(a1225 Trin. Hom.47)

That day they NegM came never out

‘That day they never came out.’ (MED)

(7) a. Ne wyrð se gitsere næfre full feos, ...

NegM becomes the miser never full of wealth

‘The miser never becomes full of wealth, …’ (Healey and Venezky (1980))

b. Ne swincke þu nefre swa muchel, ...

(a1225 Lamb.Hom.7)

NegM labour you never so much

‘You never labour so much, …’ (MED)

(8) For ði ðe he næfre nane synne ne geworhte.

(ÆCHom 11,3 21.91)

because he never no sin NegM committed

‘because he never committed any sin.’ (Healey and Venezky (1980))

OE and ME never can be viewed as an Adv Phrase, an XP category. Thus these facts indicate that OE and ME not falls under an XP category.6

6 Another argument for the XP status of ME not might be presented. It concerns examples like (i), which are among the so-called verb-second facts. As for OE not, at this point it seems to me that there are no cases such as (i) but
3. The Structure of Negative Sentences

Before discussing the ModE status of not, it will be in order here to make clear the structure of OE and ME negative sentences.

3.1. The NegP Analysis

I will begin by reconsidering the patterns (1a) and (1b). Their remarkable characteristic is that the finite verb is placed between the two negative markers and is adjacent to them. It should be noticed here that this pattern is the same as the French typical negative pattern in (9).

(9) Je ne sais pas.
    I NegM know not
    ‘I do not know.’

Besides this, the development into (1b) in ME is closely similar in process to the development into (9) in French. As for not in (1a) and (1b), it has been argued in the literature that it originally had the function of reinforcing the negative marker ne and thus it was not a negative proper, but in the ME period the typical form (1b) was established (cf. Jespersen (1917: 9), Jack (1978: 306), and Schwegler (1983: 314)). With respect to pas in (9) as well, it originally functioned as a negative emphasizer, as Schwegler (1983: 304-305) points out. He further observes that it was not a negative proper originally, but in the fifteenth century the pattern (9) became the norm.

From these considerations, it will be reasonable to assume the same structure for OE/ME and French negative sentences. As for French negative markers, the NegP analysis is proposed in studies such as translations from Latin texts.

(i) Nowt had þis doctour mynde þat he ded soo, ....
    (c1450 Capgr.St.Aug.7.17)
    Not had this doctor memory that he did so
    ‘This doctor did not remember that he had done so, ...’
    (MED)

In (i), the negative Nowt is in the topic position. It is well-known that the topic elements in verb-second constructions are generally XP categories (cf. Van Kemenade (1987), Weerman (1989), Roberts (1993), etc.). From these facts, we can say that ME not can behave as a topic like other XP categories; hence not is an XP category. Based on Roberts’ (1993) analysis, we can represent the structure of the verb-second construction with not in the topic position, as in (ii).

(ii) [CP not [C [c [finite V]]] [AGRP ... t; ...]]

In this structure, not is placed in the XP position [Spec, CP].

(10) NegP
    pas Neg’
    Neg0 ne

In (10) *pas* occupies [Spec, NegP] and *ne* the head Neg0. Following Belletti (1990) and Roberts (1993), I assume the following D-structure of French negative sentences:

(11) AgrP
    NP Agr’
    Je Agr NegP
    pas Neg’
    Neg0 TP ne T VP
    V sais

If (11) is correct, the same structure may be given to (1a) and (1b). (12a) and (12b) are the structures of (1a) and (1b), respectively.⁷

(12)

a. AgrP
   NP Agr’
   ic Agr NegP
   nawiht naht Neg’
   Neg0 TP ne T VP
   V secge

b. AgrP
   NP Agr’
   I Agr NegP
   not Neg’
   Neg0 TP ne T VP
   V seye

⁷ Note that each Infl category is head-initial in (12a). Here I assume with Pintzuk (1991, 1993) that OE allows such an Infl-medial structure as well as an Infl-final structure.
In the next section, I will show that OE and ME *ne* is an \( X^0 \) category.

### 3.2. OE and ME *ne* as an \( X^0 \) Category

Jespersen (1917: 12) and Mitchell (1985a: 661) observe that OE and ME *ne* is generally positioned immediately before finite verbs, as shown in (3)–(5). It should be noted that in this regard OE and ME *ne* behaves like French *ne* as in (9). In view of cases such as (9), Pollock (1989: 414) takes French *ne* to be a clitic. In fact, Van Kemenade (1985: 79) and Traugott (1992: 268) view OE *ne* as a clitic, and Fischer (1992: 281) also deals with ME *ne* as such. The clitic nature of OE and ME *ne* is obvious especially from the OE facts (13) and the late ME facts (14), where *ne* loses its vowel and is fused with the finite verbs.

(13) a. ne + is → nis  
    b. ne + was → nes  
    c. ne + willan → nillan  
    d. ne + agan → nagan

(14) a. ne + am → nam  
    b. ne + was → nas  
    c. ne + will(e) → nill(e)  
    d. ne + art → nart

(cf. Jespersen (1917: 12))

Its clitic nature is also indicated by verb-second facts like (15).

(15) a. ðonne ne gelyfð he nanes soðes (Bo 107.10)  
    then NegM believes he no truth  
    ‘then he does not believe any truth’

    (Mitchell (1985b: 972))

    b. Of al his strengðe ne drede we nawiht.  
    (c1225 SWard 18,168)  
    ‘We do not dread all his strength.’  
    (MED)

The existence of these cases suggests that the form ‘*ne* + finite verb’ can move to the second position in the clauses. This amounts to saying that the finite verb on which *ne* is cliticized can move to \( C^0 \), as indicated in (16), given the *Barriers*-style phrase structure (Chomsky (1986)).

(16) \[ \text{CP XP} \left[ C \left[ \text{[finite V]}_i \right] \left[ \text{AGRP} \ldots t_i \ldots \right] \right] \]

Here, following Chomsky's (1986) Structure Preserving Hypothesis for adjunction, I will assume that cliticization is a process of adjunction. Under this hypothesis, the category adjoined to a head must be

---

8 It should be noted that the form ‘*ne* + finite verb’ frequently occupies the initial position of a sentence in OE. For the analysis of this case, see fn. 9.
a head. Then it follows that *ne* is a head since it is adjoined to the verb, which is a head.\(^9\) From the discussion above, it can be justified that *ne* is generated under Neg\(^0\). This makes it clear that the structures (12) are motivated.

### 3.3. The Derivation of Negative Sentences

It remains to be examined whether the OE and ME negative sentences discussed above can be derived from (12) appropriately. Specifically, we will focus our attention on the derivations of the examples (3)-(5). Let us begin by considering the examples (3), which exhibit the patterns illustrated by (1a) and (1b). To derive (3), we will take V-movement and the cliticization of *ne* into account. As for the former, it is reasonable to say that V-to-Agr raising takes place in (3), because Agr in OE and ME was strong enough to attract a verb, as Pollock (1989: 419-420) suggests. But what is problematic for this V-movement is that the complex \([T \ V \ T]\), which has been formed by V-to-T, raises across Neg\(^0\) to Agr, as shown in (17).

\[(17)\]
\[
\text{AgrP} \\
\text{NP} \quad \text{Agr'} \\
\text{Agr} \quad \text{NegP} \\
\text{T} \quad \text{Agr} \quad \text{not} \quad \text{Neg'} \\
\text{V} \quad \text{T} \quad \text{Neg}^0 \quad \text{TP} \\
\text{te} \quad \text{T} \quad \text{VP} \\
\text{tv} \quad \text{V} \\
\]

\(^9\) Returning to the examples (4), however, one might object that *ne* is an XP category because these examples could be regarded as verb-second constructions where *ne* is a topic and the finite verbs are in the second position. But the clitic nature of *ne* is clear from (i).

(i) a. *Nis hit naht feor ascaden from ðisse worulde,...*  
\(\text{NegM+is it not far separated from this world}\)  
\('\text{It is not far separated from this world.}'\)  
\((\text{CP 51.399.26})\)  
\((\text{Healey and Venezky (1980)})\)

b. *Nis heo nost icome. (a1300 Floris (Vit) 361)*  
\(\text{NegM+is she not come}\)  
\('\text{She has not come.}'\)  
\((\text{MED})\)

Thus it is possible to say that *ne* as in (4) is also a head and again ‘*ne*+finite V’ moves to C\(^0\).
From the viewpoint of Rizzi's (1990) Relativized Minimality, this head movement should be excluded, because Neg⁰ is a typical potential antecedent governor for the trace of [T V T]. To solve this problem, following Roberts (1993), I assume (18) and (19).

(18) Relativized Minimality: in the configuration ..X..[Y ..W.. Z..], X antecedent-governs Z only if there is no W such that:

(i) W is a typical potential antecedent-governor for Z;
(ii) W intervenes between X and Z. (Roberts (1993: 32))

(19) W is a typical potential antecedent-governor for Z =

(i) ... in an A'-chain: for Z = XP, W is an A'-specifier c-commanding Z.
   for Z = a head W is an A'-head c-commanding Z.

(ii) ... in an A-chain: for Z = XP, W is an A-specifier c-commanding Z.
   for Z = a head W is an A-head c-commanding Z.

(op.cit.: 40)

What should be noticed here is that (19) contains two types of heads: A-head and A'-head. In Roberts’ (1993: 39) system, A-heads have properties that are relevant for the determination of argument structure; while A'-heads are relevant for the A'-system. More precisely, A-heads license A-chains by assigning Case or a θ-role, while A'-heads license A'-chains, e.g., by licensing an operator. Specifically, Agr assigns nominative Case within Roberts’ framework and V is in general a θ-role assigner. Thus they are relevant for the A-system and hence are considered A-heads. On the other hand, C can license an operator (e.g., wh-operator) (cf. Roberts (1993: 45, 68)) and Neg is construable as an operator position (cf. Roberts (1993: 40)). They are relevant for the A'-system and hence are A'-heads.¹⁰ Under this analysis, even if the complex [T V T] raises across the head Neg⁰ to Agr, it can antecedent-govern its trace. The reason is that the raising of the complex to Agr forms an A-chain but Neg⁰ is an A'-head and thus Neg⁰ is not a typical potential antecedent governor for the trace of the complex. As for the cliticization of ne, I assume that it raises to

¹⁰ Roberts (1993: 41) also assumes that morphologically selected heads are classified as A-heads. In his framework, Y⁰ is morphologically selected by X⁰ if X⁰ has the subcategorization frame [+ _Y⁰] which triggers substitution of Y⁰. As for T, Roberts (1993: 41) makes the following assumption: if morphologically selected, T is an A-head, and if not, it is an A'-head.
the Agr position, and it is adjoined to the verb in the complex $[\text{AGR} [T V T]]$. On this account, the derivation of (3) can be diagramed as in (20).

(20)

```
AgrP
  NP
  Agr'
Agr NegP
  T
  Agr not
  Neg'
V T Neg0 TP
nei V tne T VP
```

On the basis of (20), we can derive the other two cases. (4) can be derived by raising to $C^0$ the complex $[\text{AGR} [T \{V \text{ne} V\} T] \text{Agr}]$ formed in (20), as schematically represented in (21).

(21) $[CP [C^0 [\text{AGR} [T \{V \text{ne} V\} T] \text{Agr}]] [\text{AgrP} \{\text{NPsubj} t\text{Agr} [\text{NegP not} t\text{ne} [TP tT [VP tV]]]]]]$

With respect to (5), note that the verbs with ne are in the clause-final position. Here, I assume with Pintzuk (1991, 1993) that OE allows both head-final Infl and head-initial Infl. The latter was exemplified in (12a). But (5) can be regarded as a case of the former, which is typical of OE embedded clauses. Thus, the structure of (5) is shown as in the following:

(22) $[CP C^0 [\text{AgrP} \{\text{NPsubj} [\text{NegP not} [TP [VP \text{NPobj} tV] tT] t\text{ne}] [\text{Agr} [T \{V \text{ne} V\} T] \text{Agr}]]]$

The discussions above lead us to conclude that the examples (3)-(5) can be derived from (12) appropriately.

4. The Change of not into an X$^0$ Category

Having established the structure of OE and ME negative sentences, we are now in a position to consider the projection of ModE not in X-bar theory. It seems to me that not, which was an XP category in OE and ME, began to change into an X$^0$ category in the sixteenth century. In what follows, this will be justified.

First, it is noteworthy that the frequency of the auxiliary do increased remarkably in the negative sentences in the sixteenth century (cf.
Ellegard (1953: 162)).

(23)  
  a. I doe not disdayne you. (Udall, *Roister Doister* 24)
  b. he do not marre them. (More, *Utopia* 22)

(Engblom (1938: 137))

Here I want to claim that this phenomenon appeared because *not* began to occur in the position of Neg⁰ as an X⁰ category. This is explained below within Roberts' (1993) framework. But before we come on to the explanation, some remarks should be made concerning his analysis of ModE examples such as (24).

(24) He left.

In this case, Agr and T lower to V in overt syntax, as indicated in (25).

(25)  
```
  NP  Agr'
  |   
  He  Agr
  T   VP
  ti  ti
  T   V
  V_i  T_i
  leave T Agr_i
```

(cf. Roberts (1993: 278))

But this operation gives rise to ECP violations, since it creates traces that are not c-commanded by their antecedents.¹¹ But at LF no ECP violation results and thus (24) is derived appropriately, because V raises to C in the mapping to LF. Specifically, he regards this LF-movement as an instance of Quantifier Raising (QR), the version of move-α which adjoins operators to positions which permit their scope to be structurally determined. The element that must be raised is T, which can be regarded as a tense operator (cf. Roberts (1993: 277, 325)). It moves to the higher position C at LF to define its scope

¹¹ The relevant portion of the ECP is the antecedent-government requirement. Roberts (1993: 76, fn. 9) gives its definition as follows:

(i) α antecedent-governs β iff: α c-commands β, α is coindexed with β, no barrier intervenes between α and β, and relativized minimality is respected.
in relation to the rest of the constituents of the sentence. Since T is adjoined to V at S-structure, the complex \([V V [T T Agr]]\) raises to C at LF. It moves through T, but skips Agr, as indicated in (26). From C, V antecedent-governs the trace ti, and ti in turn antecedent-governs tj. Furthermore, tj antecedent-governs t'j.

(26)

\[
\begin{array}{c}
C' \\
C \rightarrow Agrp \\
C \rightarrow V \rightarrow j \\
NP \rightarrow Agr' \\
V \rightarrow T \rightarrow He \\
Agr \rightarrow TP \\
leave T \rightarrow Agr_i \\
T \rightarrow ti \rightarrow T \rightarrow VP \\
t_j \rightarrow V \\
t'_j
\end{array}
\]

(cf. Roberts (1993: 278))

We are now ready to discuss the relation between the status of not and the occurrence of do. As for the sixteenth-century situation, Roberts observes that V-to-Agr raising of main verbs began to disappear, and Agr-to-V lowering began to take place. He further remarks that this change began in the period 1550-1575 (cf. Roberts (1993: 249)). It can be said that the processes Agr-to-V and V-to-Agr in overt syntax coexisted around this time. Now suppose that not was still an XP category generated in [Spec, NegP] in this period, just as in OE and ME. Then it is expected that examples such as (27) were the only possible negative declarative sentences.

12 Based on Roberts' idea of T-movement, we can say that \([T \text{ do}]\), which is formed by do-insertion (cf. (28)), and the complex \([Agr [T V T] Agr]\), which is formed by overt V-to-Agr raising, undergo the LF raising to C. Besides T, Roberts (1993: 40, 338) views Neg as an X\(^0\) category which undergoes QR as an operator at LF. Roberts (1993: 338, fn.21) assumes that QR of Neg takes place for reasons connected to determining its scope domain.

13 If the complex \([V V [T T Agr]]\) moved across T to C, a violation of Relativized Minimality (18) would be yielded. This movement would form an A'-'chain, since C is an A'-'head. T is also an A'-'head, and thus T would be a typical potential antecedent governor for the trace of the complex. As for the skip of Agr, it can be said that the movement of the complex to C is a case of QR at LF and as such it skips the A-position Agr. Note that this does not give rise to a violation of (18), since Agr is an A-head and is not a typical potential antecedent governor for the trace of the complex.
(27)  a. I not say.  (= (1c'))
    b. I say not.  (= (1c))

In (27a), overt Agr-to-V lowering and LF raising of the complex \([V V \{T \ T \ Agr\}]\) are applied just as in (24). Not in (27a) is XP and hence is not a typical potential antecedent governor for the trace of the \(X^0\) category \([V V \{T \ T \ Agr\}]\) at LF, from the viewpoint of (19). (27b) is derived by overt V-to-Agr like the case of (20). Again, not is not a typical potential antecedent governor for the trace of \([_AGR \{T \ V \ T\}]\).

But in fact, not only (27) but also (23) existed in that period (cf. (1c'), (1c), and (1d)).14, 15, 16

Suppose, then, that not began to be reanalyzed as an \(X^0\) category and generated in Neg\(^0\) in the sixteenth century, although XP not also existed. Then the existence of (23) can be predicted correctly. Not in Neg\(^0\) blocks QR of the complex \([V V \{T \ T \ Agr\}]\) to C at LF. The raising of the complex to C forms an \(A'\)-chain and Neg\(^0\) is an \(A'\)-head. Thus Neg\(^0\) is a typical potential antecedent governor for the trace of the complex; hence the ECP is violated. To avoid this, do is inserted

14 Actual examples relevant to (27a) and (27b) are presented as in (i) and (ii), respectively:

(i) whose all not equals Edwards moiety.  
   (1594 Shakesp., Rich.III.I.i,250)  
   (Visser (1969: 1533))

(ii) What was his name?—We asked not.  
   (1553 Udall, Royster D.(Arb.) 37)  
   (ibid.: 1535)

15 In the literature, it has been observed that examples like (27a) were also present in the fifteenth century (cf. (1c')). As shown in (i).

(i) whiche I not presuppose.  
   (1471 Caxton, Recuyell I. 208)
   'which I do not presuppose.'  
   (Visser (1969: 1533))

Under the present analysis, V-to-I raising took place in this period. Then how can cases like (i) be derived? Let us first draw attention to Roberts' (1993: 304) observation that all the fifteenth-century examples of 'not + V' order given in Visser (1969: 1533 (§1440)) have either subject gaps or pronominal subjects. Roberts points out that this characteristic is similar to that of Stylistic-Fronting of adverbials or participles, found in Icelandic and Faroese. Following his analysis, I treat the fifteenth-century examples as cases of Stylistic-Fronting of not: cases where not is moved to a position preceding verbs which are raised to Agr. This analysis is supported by Ukaji's (1993) observation. He notes that there was a growing tendency to place adverbs such as always, ever, never, and probably immediately before the finite main verb in the fifteenth century. Ukaji further points out that the placement of not before the verb began roughly at the same time as the shift of the adverbial position, and he suggests the influence of the latter upon the former.

16 Here I regard not in (27b) (= (1c)) from 1400 to the sixteenth century as an XP category. For its grammatical status after this period, see fn. 21.
in T and raises to Agr, forming an A-chain, as shown in (28).17

\[
(28) \quad \text{AgrP} \\
\quad \text{NP Agr'} \\
\quad \text{Agr NegP} \\
\quad \text{T Agr Neg'} \\
\quad \text{do_i Neg^0 TP} \\
\quad \text{not T VP} \\
\quad \text{t_i V}
\]

Neg\(^0\) does not block this raising, since the A' head Neg\(^0\) is not a typical potential antecedent governor for the trace of \([T \text{ do}]\). Therefore, this trace is governed by \([T \text{ do}]\) in Agr; hence the antecedent-government requirement of the ECP is satisfied. These considerations lead us to

17 Unlike the case of \textit{not}, the presence of \textit{never} in present-day English does not trigger do-insertion.

(i) He never runs.

Thus it can be regarded as an XP category. With respect to its position, there are two possibilities: VP-adjoined position and [Spec, NegP]. It seems reasonable to assume that \textit{never} occupies the former position, as shown in (ii), although OE and ME \textit{never} and its variants could occur in the latter position (cf. §2).

(ii) \([\text{AgrP He Agr [TP T [VP never [VP runs]]]}]\)

This is motivated by the following perfect infinitive facts:

(iii) a. My mother would have never dreamed I would come to this lonely period. \quad \text{(Konishi, ed. (1989: 1197))}

b. He claims to have never seen her before. \quad \text{(Curme (1931: 134))}

In perfect infinitives, \textit{never} can occur between \textit{have} and a past participle. If \textit{never} were generated in [Spec, NegP], only structures such as (iv) could be given. Here I assume with Roberts (1993: 259, 316, 325) that modals and infinitive to are inserted in T.

(iv) a. \([\text{AgrP My mother [Agr wouldi] [NegP never [Neg e] [TP t_i [VP have [VP dreamed]]]]}]\)

b. \([\text{... [NegP never [Neg e] [TP [T to] [VP have [VP seen ...]]]]}]\)

As is clear from (iv), the positions of \textit{never} in (iii) cannot be captured. Rather, if \textit{never} is generated in VP-adjoined position, we can capture their positions in (iii) appropriately, as represented in (v).

(v) a. \([\text{AgrP My mother [AgrP [Agr wouldi]] [TP t_i [VP have [VP never [VP dreamed]]]]}]\)

b. \([\text{... [TP [T to] [VP have [VP never [VP seen ...]]]]}]\)

Hence, we conclude that \textit{never} in cases like (i) is a VP-adverb.
conclude that *not began to change into an X$^0$ category in that period, although XP *not also existed.

Next, I will provide indirect support for this conclusion. Let us now pay attention to the advent of *n't, the contracted form of *not (cf. (1e)). Jespersen (1917: 117, 1940: 429) observes that *n't came into use in speech about the year 1600, and in writing about 1660. In what follows, it will be shown that *n't is an X$^0$ category. Then, the fact that *n't did not appear in OE and ME but after the sixteenth century will suggest the change of *not into an X$^0$ category in the sixteenth century.

To motivate the X$^0$ status of *n't, let us turn to the analysis proposed by Zwicky and Pullum (1983), and Zanuttini (1990, 1991). Their common claim is that *n't functions as a bound morpheme, which needs to be supported by a lexical host in syntax. Here, I will present two of Zwicky and Pullum’s arguments for such status of *n't. One concerns the contrast of an attachment of *n't as in (29) with auxiliary contraction as in She's gone.

(29) *I'dn't be doing this unless I had to.

(Zwicky and Pullum (1983: 507))

In this case, I'd is first formed from I would by applying auxiliary contraction, and then *n't is attached to I'd. According to Zwicky and Pullum (1983: 504), auxiliary contraction is applied after syntax. If so, the attachment of *n't also should take place after syntax. But the ungrammaticality of (29) indicates that *n't must not be attached after syntax but rather in syntax like verbal inflections, which are bound morphemes.

The other argument is that *n't can undergo a syntactic operation together with the verb to which it is attached, as shown in (30).

(30) a. You haven’t been here.
   b. Haven’t you been here?

(Zwicky and Pullum (1983: 506))

In (30b), *n't, combined with have, is moved to the initial position. This is parallel to the syntactic movement of the combination ‘verb+inflectional affix’. For example, has ‘have+inflectional affix’ in (31a) undergoes raising and (31b) is derived.

(31) a. He has been here.
   b. Has he been here?

These facts suggest that *n’t functions as a bound morpheme which is attached to auxiliaries in syntax.

With this nature of *n’t in mind, let us consider its projection in X-bar
theory. It should be noted here that the nature of *n't* is similar to that of the clitic *ne*, which is also a bound morpheme. Recall that the combined form ‘*ne* + verb’ undergoes movement, as shown in (16). Hence, again following Chomsky’s Structure Preserving Hypothesis for adjunction, I will assume that the attachment of *n't* is a process of adjunction. Under this hypothesis, *n't* adjoined to an auxiliary verb must be an $X^0$ category, since the auxiliary verb is an $X^0$ category.\(^{18}\) Thus the advent of *not* as an $X^0$ category in the sixteenth century is indirectly supported by the fact that *n't*, an $X^0$ category, appeared only after this period.

The claims so far can be summarized as follows: In OE and ME, *not* was generated in [Spec, NegP] as an XP category, but in the sixteenth century, it began to be reanalyzed as an $X^0$ category and generated in Neg.\(^{19, 20}\) In the next section, we will consider why the reanalysis took place.

\(^{18}\) Based on Zwicky and Pullum’s (1983) claim, Kayne (1989: 9, 17) and Zanuttini (1990: 8, 1991: 90) also analyze *n't* as a functional head Neg.$^0$.

\(^{19}\) In present-day English, negative sentences with auxiliary *do* such as (1d) are normal. Thus the $X^0$ status of *not* can be viewed as being established. A question then may arise of how the inner-island effects, treated by Rizzi (1990), can be accounted for. In (ia), for instance, the clefted adverbial *for this reason* can be construed either with the matrix clause or the embedded one, while in (ib) it can only be construed with the matrix clause.

\[(i)\]
\[
\quad a. \text{ It is for this reason that I believe that John was fired.} \\
\quad b. \text{ It is for this reason that I don’t believe that John was fired.} \\
\]

(Rizzi (1990: 15))

In (ib) the construal of the clefted adverb with the embedded clause is blocked by the presence of negation. In his approach, *not* is an XP category and occupies an A’ specifier position. Under Relativized Minimality, *not* in (ib) blocks LF A’-movement of the adverbial operator.

Under the present analysis, I follow Ouhalla (1990: 220) in assuming that [Spec, NegP] contains an empty operator in present-day English.

\[(ii)\] [Neg Op [Neg’ [Neg not]]]

Given the structure in (ii), we can provide the following account: the blocking effect which negation has on the movement of the adverbial operator in (ib) is induced by the operator in [Spec, NegP].

\(^{20}\) Furthermore, it is interesting to note that verb-second phenomena discussed in fn. 6 are not found after the sixteenth century. For instance, Engblom (1938: 142–144) makes elaborate investigations into the pattern ‘negative expression + do + subject ...’ in Milton’s works. But no cases of the pattern ‘*Not* + do + subject ...’ can be found there. This indicates that *not* changed into $X^0$ category to a considerable extent after the sixteenth century.
5. The Reanalysis of not

5.1. Neg⁰ without Lexical Content

If not began to be generated in Neg⁰ in the sixteenth century, a question arises of what happened to ne, which occupied this position in OE and ME. As is clear from (1c), ne was seldom generated around 1400. As for the typical ME pattern illustrated by (1b), ne was pronounced with so little stress that it was apt to disappear altogether, as Jespersen (1940: 427) points out. Furthermore, even in OE, there were negative sentences with ne missing, as Mitchell (1985a: 672–673) points out. It seems to me that these facts are closely related to the reanalysis of not and its generation in Neg⁰. Let us now represent Neg⁰ in the negative sentences without ne as [Neg e], which has neither lexical content nor clitic nature. From the facts above, then, we can say that OE and ME allowed [Neg e]. If there is a licensing condition on [Neg e] in the grammar, we can also say that it was satisfied in OE and ME. In view of this, I argue that the condition was violated in the sixteenth century, and to avoid this, not came to fill in the position [Neg e] as a head. Then, it will be necessary to explore the condition.

5.2. A Licensing Condition on [Neg e]

In this section, we will propose a licensing condition on [Neg e]. Before entering into the discussion of this, it will be in order here to consider other empty head positions that must be licensed. For the present, we will focus our attention on the licensing of [COMP e], which is discussed by Stowell (1981), Kayne (1981), Aoun (1985), and Aoun et al. (1987). Among these studies, Stowell proposes that [COMP e] in tensed embedded clauses is licensed if it is governed by a matrix verb. Under this proposal, [COMP e] in (32a) is licensed, because it is governed by the matrix verb knew. By contrast, in (32b, c) the embedded clauses are not in the position which is governed by the verbs: hence [COMP e] is not licensed.

(32) a. Ben knew [❙ [e] [the teacher was lying]]
   (Stowell (1981: 396))

   b. *[❙ [e] [the teacher was lying]] was hardly obvious. (ibid.)

   c. *It surprises me [❙ [e] [you have heard about Roger]]
   (Stowell (1981: 397))

By analogy, I assume here that [Neg e] is also licensed if it is governed by a particular head. Based on (17), we indicate the
structure of OE and ME negative sentences without *ne* as in (33).

(33) \[ \text{AgrP} \]
    \[ \text{NP} \quad \text{Agr'} \]
    \[ \text{Agr} \quad \text{NegP} \]
    \[ \text{T} \quad \text{Agr} \quad \text{not} \quad \text{Neg'} \]
    \[ \text{T} \quad \text{V} \quad \text{Neg}^0 \quad \text{TP} \]
    \[ \text{e} \quad \text{T} \quad \text{VP} \]
    \[ \text{tT} \quad \text{V} \]
    \[ \text{tv} \]

In this structure, the possible candidates for licensers of [Neg e] are Agr, T, and V. To begin with, I examine whether or not Agr can govern [Neg e]. Here I adopt Chomsky’s (1986) definitions of government and exclusion.

(34) \[ \alpha \text{ governs } \beta \text{ iff } \alpha \text{ m-commands } \beta \text{ and there is no } \gamma, \gamma \text{ a barrier for } \beta, \text{ such that } \gamma \text{ excludes } \alpha. \] (Chomsky (1986: 9))

(35) \[ \alpha \text{ excludes } \beta \text{ if no segment of } \alpha \text{ dominates } \beta. \] (ibid.)

Agr in (33) can m-command [Neg e], because the first maximal projection that dominates Agr is AgrP, and AgrP also dominates [Neg e]. But NegP may be a candidate for \( \gamma \) in (34), because it does not dominate Agr and excludes Agr under the definition (35). However, note that [Neg e] and NegP are the same kind of projections. Belletti and Rizzi (1981), Lobeck (1987), Aoun et al. (1987) assume that if an XP category is governed, its head is also governed. Given this, it follows that Agr can govern NegP and hence its head [Neg e]. On the assumption that the licensing governor of [Neg e] is Agr, I propose the following licensing condition:

(36) \[ [\text{Neg e}] \text{ must be governed by strong Agr.} \]

I also assume that it is applied at S-structure. In terms of (36), the reanalysis of *not* can be accounted for as follows. Essentially following Pollock’s (1989: 419–420) and Roberts’ (1985: 43) views, I assume that weak Agr or Agr ‘opaque’ to \( \theta \)-role assignment appeared in the sixteenth century, because most of verbal inflections were lost. Agr in OE and ME was strong and hence could satisfy (36). But the weak Agr in the sixteenth century could not. Therefore, to avoid this, *not* came to fill in [Neg e]: it came to be reanalyzed as an X\(^0\) category generated in Neg\(^0\).
Next, let us consider T and V. They also can govern \([\text{Neg e}]\), since they m-command \([\text{Neg e}]\): the first maximal projection that dominates them is AgrP, and AgrP also dominates \([\text{Neg e}]\). If they are regarded as licensing governors of \([\text{Neg e}]\), the conditions (37) might be assumed.

\[
\begin{align*}
\text{(37) a. } & [\text{Neg e}] \text{ must be governed by } T. \\
\text{b. } & [\text{Neg e}] \text{ must be governed by } V.
\end{align*}
\]

However, if the auxiliary verbs have and be are taken into account, a problem arises for these conditions: we cannot account for the fact that \([\text{Neg e}]\) ceased to be licensed in the sixteenth century. Note that such auxiliary verbs continued to raise to Agr after the sixteenth century, as shown in the tensed sentences they have not left/they are not students. Here the complexes \([T [V \text{ have}] T]\) and \([T [V \text{ are}] T]\) are moved to the left of not, which is assumed to be in [Spec, NegP] under the present analysis. This gives rise to an undesirable result: after the sixteenth century, \([\text{Neg e}]\) was still governed and licensed by T and V, and hence did not need to be filled with not.

Another problem will arise even if the candidates for V in (37b) are limited to main verbs, which were not raised after the sixteenth century. Here I tentatively present this idea as in (38) for expository convenience.

\[
\begin{align*}
\text{(38) } & [\text{Neg e}] \text{ must be governed by a main verb.}
\end{align*}
\]

(38) incorrectly expects that OE and ME equivalents of the auxiliary verbs have and be, which were raised to Agr, could not license \([\text{Neg e}]\). Therefore, (37) and (38) are untenable: T and V cannot be viewed as plausible licensing governors.

The considerations above make it clear that (36) is an appropriate licensing condition on \([\text{Neg e}]\). This leads to the following conclusion:

\[
\begin{align*}
\text{(39) } & \text{In the sixteenth century, the condition (36) began to be unsatisfied because of the advent of weak Agr. To avoid this, the reanalysis of not took place: not came to fill in the position } [\text{Neg e}] \text{ as a head.} \quad 21
\end{align*}
\]

---

21 One might point out that the late ME pattern illustrated by (1c) (i.e., I say not.), where not is an XP category, survived after the sixteenth century, although Agr was weak. But according to Visser’s (1969: 1534) and Jespersen’s (1940: 428) observations, it is only a small number of verbs (e.g., know, care, doubt) that continued to be used in this pattern after the sixteenth century. Roberts (1993: 253) treats ModE cases of this pattern as fixed expressions. On the basis of this view, we can say that these cases are relics of the late ME period, and not maintains its XP status only in such fixed expressions.
5.3. Facts from Other Languages

In this section, we will examine whether or not (36) also works appropriately in other languages. If (36) is adequate, correct consequences should follow in other languages as well. Here, we will focus on French, Piedmontese, and West Flemish.

Let us first examine French facts. The negative marker *ne* is deletable, as shown in (40) (cf. Ashby (1981), Schwegler (1983)).

(40) Je (ne) sais pas ...

‘I don’t know ...’  (Ashby (1981: 678))

According to Chomsky’s (1991) and Pollock’s (1989) analyses, Agr in French is strong. Based on (11), I show the relevant structure as in (41).

(41)

```
AgrP
  NP
    Agr'
      Je
      Agr
      NegP
        T
        Agr
        pas
        Neg'
          V
          T
          Neg0
            TP
              sais
                e
                  T
                  VP
                      tT
                        V
```

Here the strong Agr governs [Neg e]; hence the fulfilment of (36).

We will next turn to Piedmontese. Zanuttini (1991: 22-31, 57) shows that the negative marker of this language *nen*, an equivalent of French *pas*, falls under an XP category and it is generated in [Spec, NegP]. Zanuttini (1990, 1991) observes that this language has no equivalent of French *ne*, which suggests that it allows [Neg e]. Kayne (1991: 652) notes that V raises to I in this language. This indicates that Agr is strong. Consequently, a derivation similar to (41) is possible, and thus [Neg e] is licensed under (36).

Finally, let us consider West Flemish. Following Haegeman and Zanuttini (1991: 238), I give the structure in (42) to negative sentences in this language:

(42) [CP [ AgrP [ NegP [ TP [ VP ... V ... ] T ] [ Neg (en) ] ] Agr ]]

The negative marker *en* is deletable, as shown in (42); hence [Neg e] is possible. In this language, V-to-C raising takes place in main clauses, and V-to-I raising takes place in subordinate clauses (cf. Haegeman and
Zanuttini (1991: 247)). This shows that Agr is strong. In (42), the strong Agr governs $[\text{Neg e}]$. Therefore (36) correctly predicts that $en$ is deletable. As is clear from the discussions above, (36) is supported by the facts from other languages as well.

5.4. An Identification Principle for $[\text{Neg e}]$

In the preceding two sections, it was shown that the reanalysis of not can be accounted for in terms of the licensing condition (36), and some correct consequences follow from it in other languages. But it seems to me that this condition is not sufficient to allow the existence of $[\text{Neg e}]$.

Here it is worth considering $[\text{Neg e}]$ in the light of the theory of other null elements. Rizzi (1986, 1990) and Roberts (1993) view the theory of variable, NP-trace, PRO, and pro as consisting of two components. One is a principle of formal licensing, which characterizes the formal environment in which the null element can occur. The other is a principle of identification, which recovers some contentive property of the null element from a phonetically realized environment. Chao (1988) attempts to extend this theory to ellipted constituents as in null complement anaphora (43a), sluicing (43b), and VP ellipsis (43c) (for the formal licensing of these constituents, see also Lobeck (1991)).

\begin{align*}
(43) & \quad \text{a. John suggested that we leave, and Bill agreed $[\text{CP e}]$}. \\
& \quad \text{b. Somebody left, but we don't know who $[\text{IP e}]$.} \\
& \quad \text{c. John left yesterday, and Bill did $[\text{VP e}]$ too}. \\
& \quad \text{(cf. Chao (1988: 77))}
\end{align*}

As for $[\text{Neg e}]$, we must not ignore the fact that it can be allowed only when negative markers such as OE and ME not occur. In other words, $ne$ cannot be deleted without such negative markers. This suggests that it is necessary to assume not only the licensing condition (36) but also some principle which requires $[\text{Neg e}]$ to receive some negative content from such negative markers.

From these assumptions, it seems to be plausible to regard the theory of $[\text{Neg e}]$ as consisting of the two components like the case of other null elements. It is possible to say that (36) was offered as a case of the principle of formal licensing. It characterizes the formal environment in which $[\text{Neg e}]$ can occur. From the viewpoint of the principle of identification, we can say that the content of $[\text{Neg e}]$ must be recovered through certain negative markers. Then a device for this requirement will be necessary.
In Rizzi's and Roberts' theory of null elements, identification is performed through a coindexing relation in some cases. For example, the feature content of NP-trace and variable is recovered through coindexing in the formation of an A- or A'-chain with an antecedent or an operator, respectively. The content of pro in subject position is recovered through coindexing Infl with pro, and thus pro has the grammatical specification of the features on Infl coindexed with it. In the spirit of this theory, I propose (44) as an identification principle for [Neg e].

(44) If X is a negative marker in [Spec, NegP], then [Neg e] has the specification of certain negative features on the X coindexed with it.

In other words, the content of [Neg e] must be recovered through coindexing a negative marker in [Spec, NegP] with it under a Spec-head relation. If the negative marker is not, the relevant representation is as follows:

(45) [NegP noti [Neg' [Neg e]]]

If not is missing, as in (46), then the content of [Neg e] cannot be recovered, and thus this case is excluded.

(46) *[NegP [Neg' [Neg e]]]

6. Conclusion

In this paper, it has been argued under the NegP analysis that not occupied the position [Spec, NegP] in OE and ME, and came to be reanalyzed as an X0 category generated under Neg0 in the sixteenth century. To account for this change, I paid attention to the fact that the negative marker ne generated under Neg0 was sometimes missing in OE and ME. Then I assumed [Neg e] for cases where ne did not appear, and proposed the formal licensing condition (36): [Neg e] must be governed by strong Agr. On the basis of this, I explained the change as follows. In OE and ME, Agr was strong, and [Neg e] was governed by the strong Agr. Hence the condition was fulfilled. But Agr became weak in the sixteenth century. Then the condition began to be violated. To avoid this, not came to fill in [Neg e]. In this way, not was reanalyzed as Neg0. I further showed that some correct consequences follow from (36) in other languages. Finally, an identification principle for [Neg e] was taken into account besides the licensing condition. I assumed that [Neg e] has the specification of certain
negative features on a negative marker coindexed with it.

REFERENCES

Jespersen, Otto (1917) *Negation in English and Other Languages*, Bianco Lunos Bogtrykkeri, København.
Kayne, Richard (1989) “Notes on English Agreement,” ms., Graduate Center,
CUNY.


Ukaji, Masatomo (1984) “Kindaieigo no Hiteikozo ni Mirareru Mittsu no Mondaiten (Three Traits in the Development of the Negative Sentence in
Modern English),” *Studies in Modern English* 1, 30-37.


Faculty of Liberal Arts
Aichi Gakuin University
12 Araike, Iwasaki-cho
Nisshin-shi, Aichi 470-01