Economy, Innovation, and Prescriptivism: From Spec to Head and Head to Head*

ELLY VAN GELDEREN

Department of English, P.O. Box 870302, Arizona State University, Tempe, AZ 85287-0302, USA; E-mail: ellyvangelderen@asu.edu

Key words: economy, grammaticalization, head, Middle English, Minimalism, negation, Old English, prescriptivism, relative, specifier

Abstract. This paper decribes cyclical changes in negative and *wh*-constructions as a change from Spec(ifier) to Head. It accounts for this change through an economy principle that says 'if possible, be a head'. The changes examined all show a tendency towards heads and head-checking but execute this in slightly different ways. In addition, innovations introduce new specifiers, and prescriptive rules retain them, counteracting the effects of economy.

Changes from Head to Head and from Spec to Spec also occur. These proceed typically towards positions higher in the tree and can be explained via a 'merge over move' economy principle. The change involving heads I'll look at is the change of to from preposition *to* complementizer, and the changes involving specifiers involve French negatives and English relatives. Thus, certain instances of grammaticalization can be accounted for in structural terms.

1. Economy, grammaticalization, phrases, and heads

As is well-known, lexical verbs become auxiliary verbs (e.g., *have* and *will*) and prepositions become complementizers (e.g., *for* and *like*) through grammaticalization. These changes often involve a lexical head that is reanalyzed as a grammatical (or functional) head. The list of "grammatical structures in order of increasing grammatical function" (Newmeyer 1998, p. 227) consists of: lexical categories, functional categories and

---

pronominal elements, clitics, derivational affixes, and inflectional affixes (see also, e.g., Heine et al. 1984; Traugott and Heine 1991; Abraham 1992, 2002). In this paper, I examine the change from head to head as well as a change that has not received as much attention, one where phrases, especially specifiers, become heads.[1] The next stage, where heads disappear, is a common occurrence but will not be a focus in this article.

Within the generative tradition (e.g., Chomsky 1986), syntactic structures are built up using general rules, such as that each phrase consists of a head (X), a complement (ZP) and a specifier (YP) as in (1):

(1)   XP
            ┌─────────┴─────────┐
           YP   X′
                          ┌─────┴─────┐
                          X   ZP
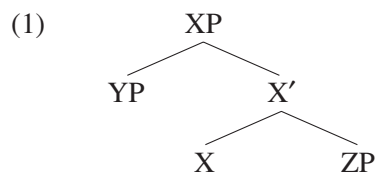
In early work, this schema is quite strict, e.g., specifiers and complements are always full phrases. With the introduction of (Minimalist) bare phrase structure in the early 1990s, this changes. A verb and a pronoun object can merge, as in (2), with one of the two heads projecting, in this case V:

(2)   VP
          ┌─────┴─────┐
          V   D
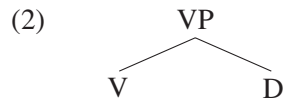
In Chomsky (1995, p. 246), it is even more extreme, and the words project, as in (3):

(3)   the
          ┌─────┴─────┐
         the   book
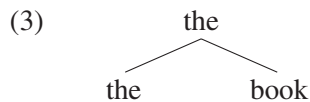
Phrase structures are built using Merge and Move. 'Merge' combines two items, e.g., *the* and *book*, of which one projects. The VP domain is usually seen as the thematic-layer, i.e., where theta-roles are determined. After functional categories such as I and C are merged to VP, Move (e.g., Chomsky 1995, p. 250) raises heads and phrases so that features can be checked in the IP and CP layers. Both head-head checking and Spec-head checking occur. In this paper, using general Minimalist principles, I argue that checking between two heads, also referred to as incorporation, is more

---

[1] The head-dependent switch is talked about in the grammaticalization literature (e.g. Haspelmath 1998, pp. 333ff.).

economical than between a specifier and a head.[2] This is part of a larger principle, as in (4):

(4)   Heads-over-Phrases
       Be a Head rather than a Phrase (if possible).

Stated as in (4), the principle holds for Merge (projection) as well as Move (checking).

Within recent Minimalism, there is a second economy principle, namely (5), (see, e.g., Chomsky 1995, p. 348). Merge, as in (2), "comes 'free' in that it is required in some form for any recursive system" (Chomsky 2001b, p. 5) and is "inescapable" (Chomsky 1995, pp. 316, 378):

(5)   Merge over Move.

Principle (5) says it is less economical to merge early and then move than to wait as long as possible with Merge. This reduces to (6):[3]

(6)   Late Merge
       Merge as late as possible.

Using (6), I will argue that if, for instance, a preposition has become less relevant to the argument structure (as in the case for *of* and *to* in English), it will tend to merge higher (in IP or CP) rather than merge early (in VP) and then move.

Both (4) and (6) are principles of Economy, which work in grammars of speakers as well as in leading children to build their grammars in a particular way.[4] However, language does not only change in the direction guided by these Economy principles. Jespersen (1921, chap. 14, section 6), in discussing the 'Ease Principle', puts it this way: "the correct inference can only be that the tendency towards ease may be at work in some cases,

---

[2] It may be that just moving features is even more economical, as Chomsky (1995, pp. 262ff.) suggests, or Agree between a probe and goal, as in later work.

   Chomsky does not discuss the Head Movement Constraint in relation to (2) and (3). As a reviewer brings up, however, the HMC is a problem for clitic movement as well as for my account of pronominal head-checking. I will just assume that whatever account works for clitic-heads works for pronominal heads as well.

[3] Chomsky (2001b, pp. 7–8) reformulates the notions of Merge and Move as external and internal merge respectively. "Argument structure is associated with external merge (base structure); everything else with internal merge (derived structure)" (p. 8). The latter leaves a copy in place but is otherwise similar to Merge. I will here observe (6) since not everything dealt with by the original Merge is relevant to theta-structure (e.g., C and I are not).

[4] There is little work on children using heads over phrases even though early sentences use N over NP.

though not in all, because there are other forces which may at times neutralize it or prove stronger than it". Lightfoot (1979, p. 384) sees borrowing and expressivity as external factors. He distinguishes (p. 405) between "changes necessitated by various principles of grammar" and those "provoked by extra-grammatical factors", hence, between necessity and chance. In this paper, I provide examples of cyclical changes of both chance (e.g., which prescriptive rule is enforced) and of necessity (e.g., Spec to Head).
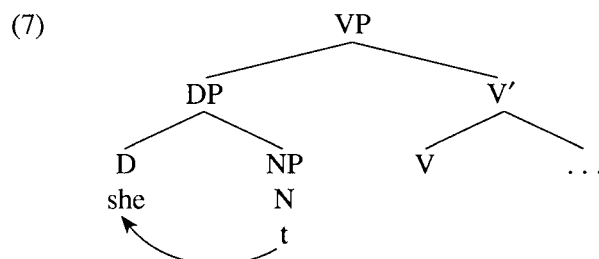
The outline is as follows. In section 2, I discuss changes in pronouns, showing that head-checking is used whenever possible over phrase-checking, in accordance with principle (4). Then, in section 3, I show how principle (4) accounts for changes affecting relative constructions, negation, and the complementizer/pronoun *whether*. In section 4, I examine head to head changes and use (6) to account for this as well as for Spec to Spec changes. Section 5 is a conclusion.

2. Subject Pronouns in Modern English

Chomsky (1995, p. 265) argues that movement carries along as much as is absolutely necessary. He does not go into the question of head versus phrasal movement but the conclusion seems obvious. In cases where it is possible to strand, stranding is obligatory. Thus, it is cheaper to move a pronoun as a head than as a phrase, and checking will be via Head to Head rather than via Spec to Head. I will show that this is indeed the case. Pronouns move as heads if they can but as phrases if they are part of a coordinate phrase or are modified by a relative clause. When phrasal, the Case is often a default one, not always clearly nominative in subject position.

Cardinaletti and Starke (1995, p. 36), following an older literature, analyze pronouns as being of three kinds: clitics are 'deficient heads', weak pronouns are 'deficient XPs', and strong pronouns are 'non-deficient XPs' (XPs being full phrases). In their discussion of, for instance, French, they argue that "the strong variant can be used only if the deficient variant is not accessible" (p. 33 bold type omitted), e.g., in the case where an adverb separates it from a verb or with coordination. The weak pronoun "remains an XP on the surface . . . , while . . . resisting coordination or modification" (p. 36). Being an XP while resisting modification seems incompatible, and I will therefore reformulate Cardinaletti and Starke's three-fold distinction as a two-fold one: elements are either X or XP, but whenever possible, XPs can be X.

A traditional analysis of pronouns is that pronouns move to D from N, as in (7):[5]

(7)

```
                    VP
             _____/  _____
           DP                 V'
        __/  \__            __/ \__
       D        NP         V      . . .
      she        N
        ↖_____ t
```

The pronoun is argued to occupy D because other Ds cannot occur together with pronouns, e.g., *the he*, *her she*, and *that I*. The reason it starts out in N is that it has some lexical features (certainly in the case of first and second person, cf., e.g., Benveniste 1966, chap. 18, see also note 7). In (7), to check the features of *she*, either the entire DP or the D moves to a position higher in the tree. Movement of the head, as in (8), will occur, I argue, if nothing else prohibits it:

(8)

```
              IP
         ____/  \____
                     I'
                __/     \__
              I            VP
           __/ \__       __/  \__
          D      I     DP        V'
         she         __/ \__    __/ \__
          ↖___      D     NP   V    . . .
              \     t      N
               _____ t
```
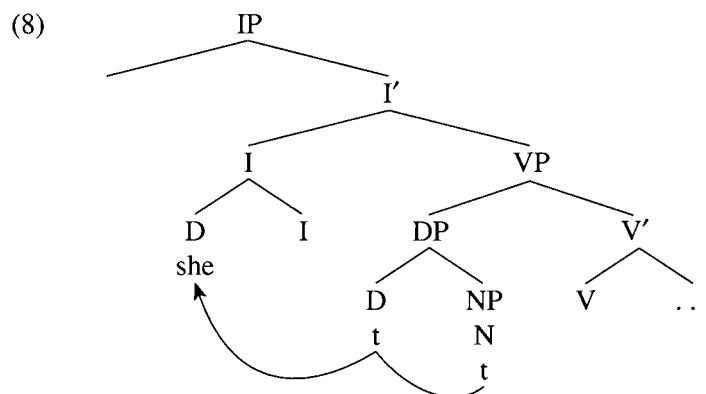
A bare phrase structure alternative for (7) would be along the lines of (2). In this case, the economy principle favoring heads over phrases, i.e. (4), does not apply to what moves but to how items are merged. For the purposes of this paper nothing hinges on which version, (2) or (7), is adopted. Full NPs, or coordinated pronouns, are different since they are forced to be phrases and check as phrases. I will now turn to the actual evidence that whenever possible, pronouns will project and move as heads.

---

[5] It could also be base generated in D, as in, e.g., Everett (1996, p. 12), and be an intransitive head (without an NP), unlike the article which would be transitive (with an NP complement).

The first piece of evidence that pronouns typically function as heads, i.e., incorporate, involves modification. For instance, they are not preceded by adjectives or modified by (restrictive) relative clauses. The reason is that they 'prefer' to be heads. So, *nice* they is not possible,[6] and relative clauses occur very infrequently.[7] In the Corpus of Spoken Professional American English (hence CSE, see note 1), the 161,000-word Faculty Meetings' part shows a split between nouns and pronouns. Thus, nouns such as *faculty*, *department(s)*, and *school(s)* are coordinated in approximately 10% of the cases, e.g., *faculty* occurs 749 times and is coordinated with *and/or* 70 times, i.e., 9.3%, and *department(s)* occurs 188 times, of which 18 are coordinated, i.e., 9.6%.[8] Nouns are hard to distinguish in terms of subject or object, and the 10% constitutes all nouns. The same texts contain 2,748 first person singular nominative pronouns, but only 15 of these are coordinated, and none are modified. This means only 0.5% are phrases. Accusative forms are slightly more frequent (the numbers are statistically significant at $p < 0.05$), but the numbers are small (6 coordinated out of 322 makes 1.55%). Second person pronouns are not distinguished for Case and hence include both subject and object pronouns. They occur 1,434 times, of which 8 are coordinated, 3 appositive, and 26 have a *wh*-relative. This means they are phrasal 2.6% of this time. This corpus has fewer third person singular nominative pronouns, namely 246 *s/he*, and of these 9 are coordinated but none modified. So, third person is a phrase in 3.7% of the cases. These figures indicate (a) that the pronoun versus noun difference is very distinct, as shown in Table 1, for all first person pronouns against two nouns. It may also indicate (b) that there is a person split since first person is the least often phrasal (statistically significant between first and second and between first and third at $p <$

---

[6] First and second person pronouns have more nominal characteristics, hence *unlucky we/us* is possible, and *we linguists*. Lyons (1999, pp. 27, 141) argues that first and second person pronouns are Ds and that the absence of third person *they linguists* is due to *the* being the third person D rather than *he/she/they*.

[7] If they occur, they are usually non-restrictive, as in (i), or if the pronoun is a generic, as in (ii). There are many archaic sounding ones ('blessed are they that . . .'):

   (i)   But if he, who had undoubtedly been always so much the most in love of the two, were to be returning with the same warmth of sentiment which he had taken away, it would be very distressing. (Jane Austen, *Emma*, Vol. 3, chap. 1)

   (ii)   He who believes in separate and innumerable acts of creation will say, that . . . (Charles Darwin, *Origin of Species*, chap. 6)

Pronouns in (i) and (ii) are forced to be phrases, but, as we'll see below, pronouns are generally not modified.

[8] Nouns that occur less often vary more, as expected. For the noun *school(s)* in CSE-FAC 1995, the percentage is 13.8, namely 29 instances of *school(s)* with four coordinated. In the same part, *students* are coordinated 7 times out of 81, which is 8.6%.

Table 1. First person pronouns versus nouns in CSE-FAC, X-square is 205.698, $p < 0.001$.

	Head		Phrase		Total
I and me	3049	(= 99.3%)	21	(= 0.7%)	3070
faculty and department(s)	88	(= 9.4%)	849	(= 90.6%)	937

0.001 but not between second and third), and hence first person more often projects as a Head.

When personal pronouns are 'pushed into' being full phrases through coordination, as in (9), the Case is often a default Case:

(9)   Huck and me ain't cry-babies.

<div align="right">(Mark Twain, Tom Sawyer, chap. 16)</div>

This is true in most colloquial usage, as (9) shows, and (12) below, and confirms that checking through Spec-Head in (9) is different from checking through Head-Head with non-coordinated pronouns.

Other, less direct, evidence that pronouns are heads is that there are varieties of English, e.g., North East Yorkshire (Cowling 1915, cited in Chapman 1995) where agreement on the verb disappears if the pronoun is adjacent, indicating that the pronoun has 'become the agreement', and that means it is a head. In these dialects, there would be a choice of either a verbal ending or an adjacent pronoun. I won't provide an analysis as to whether the Verb moves to I in these dialects or whether the pronoun+V combination is listed as such in the lexicon. Chapman (1995) divides English dialects of the British Isles into three groups: those in southwest England and East Anglia where no inflection is left; those in south east England where the *-s* ending is used throughout the paradigm; and those of the 'North Country', where "the use of *-s* as a marker of all persons singular and plural in the present is determined by the type of subject with which the verb agrees" (Chapman 1995, p. 36). If a pronoun is immediately adjacent to the verb, as in (10), there is no *-s*; if it isn't, the *-s* occurs, as in (11):

(10)   I tell him not to

(11)   I often tells him
         (both from Chapman 1995)

This suggests that the pronoun forms the agreement.[9] It also shows that pronouns need not incorporate if, for instance, they are coordinated:

---

[9] In earlier English (see Newmeyer 1998, pp. 270–271), the second person pronoun cliticizes

(12)   him and me drinks nought but water
        (also from Chapman)

Another tendency in this variety is that only the pronouns that are adjacent to the verb get nominative Case, as in (10). This trend is not absolute, as (11) shows. Wolfram et al. (1999, p. 70) show how these tendencies are reflected in American dialects. For instance, agreement tends to be lacking with pronominal subjects.

Evidence that pronouns are heads from code-switching involves subjects and finite verbs. For many language pairs (e.g., Jake 1994, Nortier 1990), a switch between a full NP subject and a verb, as in (13) between English and Dutch, results in an acceptable switch, but in the case of a pronoun and verb, as in (14) again between English and Dutch, it does not:

(13)   Those awful neighbors schijnen altijd   herrie te
        Those awful neighbors seem   always noise to

        moeten   maken.
        have to   make

(14)   *They schijnen altijd   herrie te moeten   maten.
        They   seem   always noise to have to   make

This is usually described as a prohibition against the switching of grammatical categories, but it can be explained, as in, e.g., MacSwan (1999), in terms of a prohibition against language mixing at the word level in the case of a pronoun and a verb. In the languages where pronouns and verbs can code switch, the pronouns are emphatic (see, e.g., Jake 1994), i.e., phrasal.

A minor piece of evidence that pronouns incorporate (i.e., move as heads) is that when pronouns are separated from the verb by an adverb as in (15) or interjection, as in (16), they are often repeated. This is

---

to the verb if the latter is in C, as in questions such as *What wiltow seyn of this* (HC-ME3). I won't go into this. There is perhaps something special with second person since in Dutch verbal inflection is reduced in exactly those circumstances.

In many of the modern Germanic languages (e.g. Dutch and Swedish), there is a restriction for many speakers not to separate subject pronouns from C, as in (i) for Dutch and (ii) for Swedish (see e.g. Holmberg 1991, p. 166):

(i)   Heeft mogelijk *ze/Monica het niet gelezen?
         Has   possibly   she   it   not   read

(ii)   Har möjligen *hon/Monica inte sett   boken?
         Has possibly   she/Monica not   seen the-book

The reason may be that these pronouns need to incorporate. I will just restrict myself to English.

reminiscent of subject doubling in, e.g., Northern Italian and colloquial French:

(15)   they apparently – they're involved in this. (CSE-WH96B)

(16)   We find they, you know – they both work comfortably, to be honest with you. (CSE-COMR6B97)

In coordination of verbs, as in (17), the pronoun is often doubled as well, unless the two verbs form a semantic unit, as in (18a), or a special subordinate, as in (18b):

(17) a. but I do. And I bet Judith does. (CSE-COMR797)
     b. Yes I do, and I am absolutely convinced . . . (CSE-FACMT97)

(18) a. And I hope and expect that we could continue to improve . . . (CSE-WH94)
     b. when I try and work with elementary teachers and try to map stories . . . (CSE-COMR6A97)

Without formulating a theory of ellipsis, it is hard to arrive at relative numbers for the two constructions. I will leave that for further research.

Lambrecht (1981) makes a strong case for Non-Standard French as a pronominal argument language (to borrow Jelinek's term) since (19a) is ungrammatical and needs to have the subject repeated, as in (19b). Even in Standard French, (19a) is marginal, and (19b) is preferred:

(19) a.*Je lis   et   écris
        *I   read   and   write

     b. Je lis   et   j'écris
        I   read   and   I-write

So (19) in French shows that pronouns are further on their way to becoming agreement markers than the pronouns in (17) and (18) are in English.

A last argument is that if pronouns check as heads, they might be more often 'cliticized'. This is indeed the case as Table 2 shows. Cliticization of *is* (or *has*) is much more common to a pronoun than to an NP[10] in the CSE 1-million word White House briefings' part:

---

[10]   The difference is statistically significant at $p < 0.001$, X-square 853.435. No other noun occurs with as much frequency in the CSE as *president*, but other nouns show the same tendency, e.g. *teacher* is followed 8 times by *is* but never by *'s*; *work* is followed 34 times by *is* and never by *'s*.

Table 2. Cliticization to pronouns vs NPs in CSE-WH.

	's		is + has	
he	1456	(= 65%)	298 + 493	(= 35%)
president	83	(= 8.6%)	358 + 522	(= 91.4%)

If pronouns check features in a head-head configuration, as in (8), they can start to be seen as agreement markers. Standard English is not that far (yet), but further along perhaps with first than with third person. Pierce's (1994) data comparing French and English child language shows that French subject pronouns, in accordance with Lambrecht's (1981) observations, are much further along to being agreement markers than English ones. French children use pronouns only with finite verbs (which in the singular have the same null form), but NPs occur with non-finite verbs. Coordinated pronouns are not used, and non-emphatic pronouns are also repeated in (20), showing they are needed for proper checking:

(20)   Moi je sautes et   je descends   (Pierce 1994, p. 329)
        me   I   jump   and   I   go-down

English children, in contrast, use pronouns with infinitives as well, coordinate pronouns, and need not repeat them in sentences such as (20).[11]

The view that agreement represents an argument, as Jelinek (1984) and Willie (1991) have argued for, e.g., Navajo, is similar to the view that the agreement on the verb in Spanish and Italian licences the (subject) argument (e.g., Safir 1985). Recent work by Taraldsen (1992) and Ordonez and Trevino (1999) argues that the (subject) NP in languages such as Spanish is an adjunct, not an argument. However, as Jelinek (2001) makes clear, all NPs are adjuncts in pronominal argument languages, and that is certainly not the case in (English and) French. In the latter, only definite NPs can be adjuncts.

In conclusion to section 2, standard English pronouns have the option to be heads (move as D) or to be phrases (move as DP). Whenever possible, heads will strand their projections and move as heads and

---

[11]   Neurolinguistic work on the difference between pronouns and lexical NPs is not decisive. Even though Kean's (1979) work would suggest that if pronouns incorporate they would be harder for Broca's aphasics, this is not borne out in work such as Friederici et al. (1991). However, a lot of the neurolinguistic research on pronouns is done on understanding co-reference, not on agreement.

incorporate into I to check Case. This does not mean pronouns themselves are the agreement markers on the verb (but there are varieties, with constructions such as (10) and (11) above, where this is the case). There is a person split where coordination and modification is concerned, indicating that first (and second) person pronouns are less often phrasal than third (perhaps caused by differences in inflection).

3. Spec to Head: 'Be a Head, if possible'

Having just shown that, whenever possible, pronouns behave like heads and not like full phrases, I now show that this economy principle also holds in language change: change will be from Spec to Head. All of the changes below show the interesting interaction, talked about in, e.g., Jespersen (1921, chap. 14, section 6) as a 'tug-of-war', between economy and innovation. Economy eliminates Specs; innovations reintroduce them (e.g., reinforcing *not* and *wh*-relatives); and prescriptive rules either stop their introduction (e.g., the ban on multiple negation) or try to stop their change to head (e.g., rendering *whether* phrasal by adding *or not*). Under this view, grammaticalization is uni-directional, caused by structural factors.

3.1. Relative Pronouns

In 3.1.1, I first sketch the Modern English situation and show that even though heads are preferred, there are (prescriptive) forces at work that favor specifiers. In 3.1.2, I outline some uncontroversial views on the history of relatives and argue (against the conventional belief) that relative *that* has a different status in Old English than it has in Modern English and that this is a case of a phrase/Spec becoming a Head. In 3.1.3, the reason for this change is again argued to be the Heads-over-Phrases Economy Principle in (4).

3.1.1. A Head preference in Modern English?
In Modern English, relatives are formed using *that* or a *wh*-pronoun in the case of restrictives, as in (21), and a *wh*-pronoun in the case of, relatively infrequent, non-restrictive relative clauses, as in (22). A *wh*-element and *that* are not allowed to occur together in standard English, as (23) shows, but occur in certain varieties of English, e.g., in Belfast-English, as in (24), as in other varieties of Germanic:

    (21)   The woman that/who lives next door is about to travel to Mull.

(22)   Clinton, who was president two terms, is thinking of running
        again.

(23)   *The woman who that I met yesterday is great.

(24)   I wonder which dish that they picked.   (Henry 1995, p. 107)

As in the case of interrogatives, the relatives are involved in checking the
features of the C (see Rizzi 1990 for an early version of the *Wh*-Criterion).
Using a (non-expanded) CP model, the *wh*-element would be in the spec-
ifier of CP and *that* in the head, and both cannot occur (in Modern English).
Assuming *that* in (23) is the head of CP accounts for the absence (again
in standard English) of (25) and (26), where of *that* and *that's* are XPs:

(25)   *The woman of that I saw a picture . . .

(26)   *The woman that's picture I saw . . .

There could in principle also be evidence in terms of extraction, but this
can't be tested since relative clauses are islands (part of the complex NP
constraint).

    In Modern English, there is a strong tendency in relative clauses to use
heads, such as *that* (and *as*), rather than specifiers, such as *wh*-pronouns.
This is shown in Table 3 for the 2-million spoken CSE. This corpus was
searched for the string [article + noun + relative-pronoun]:

Table 3. That versus who.

	the N		a(n) N	
that	5637	= 82%	1758	= 81%
wh-form	1199	= 18%	414	= 19%

Others have noticed similar trends, e.g., Romaine (1982).

    There are several reasons that the change to heads has not been
completed. Many prescriptive rules concern relatives, and these favor *wh*-
relatives over *that*, for instance, the rule that *who* is to be used for humans.
Fowler (1926 [1950], p. 716) says "at present there is much more reluc-
tance to apply *that* to a person than to a thing. Politeness plays a great
part", and a more recent guide says: "*who* refers to people or to animals
that have names. *Which* and *that* usually refer to objects, events, or animals
and sometimes to groups of people" (Kirszner and Mandell 1992, p. 381).
There is a second rule that favors *wh*-relative, i.e., specifiers, namely, the
rule against preposition stranding. Sentences that end with prepositions,
such as (27), are judged to be incorrect, and (28) is preferred:

(27)   I met the woman who I had seen a picture of.

(28)   I met the woman of whom I had seen a picture.

Since (28) is only possible with *wh*-relatives, this rule again favors *wh*-relatives.

It is also interesting to notice that very few *wh*-elements are Case-marked. For instance, in the CSE, there is one instance of a noun followed by *whom*, as in (29), but hundreds with *who*. So perhaps *who* is becoming a head, not expressing Case (only phrasal categories have the functional categories present to be able to have Case):

(29)   and that the president, whom I think you've all heard on this subject, is – he has . . . (CSE-WH97B)

This occurrence of *who* without Case marking may be the reason it is a head 'competitor' to *that*. There are Norwegian dialects (e.g., Taraldsen 1985) in which the *wh*-element when it is monosyllabic does not bring about Verb-second, as in (30), but does when it is a phrase, as in (31). This is explained if the *wh*-element has changed from Spec to Head blocking Verb-movement, in accordance with the Economy Principle:

(30)   Ka   dokker sa
        what   you   say

(31)   Ka   for nokka   sa   dokker
        what for something   say   you
                                    (both from Taraldsen 1985, p. 21)

3.1.2. *Changes from Spec to Head*

Turning now to the historical data, I show that there are innovative cyclical changes where specifiers become heads, in accordance with (4), but where new specifiers are introduced. For Old English, it is widely accepted in the literature that the regular relative complementizer is *þe* (e.g. Quirk and Wrenn 1955, p. 72; Allen 1977). Another relative clause marker is the demonstrative pronoun (*se*, *þat*, etc.) with *þe*, as in (32) or without *þe*:

(32)   Æghwæþres sceal scearp scyldwiga   gescad   witan,
        every   shall sharp   shield-fighter difference know,

        se   þe wel þenceþ,   worda ond worca,
        the that well thinks/judges-S words and works

        'Every sharp shield fighter, who judges well, must know the difference between words and works'. (*Beowulf* 287–9)

There is evidence that the demonstrative originally belongs to the main clause since the case in OE is often that of the main clause, unlike in Modern English. Hock (1991, p. 342) mentions that "similar patterns are found in Old Norse and traces also in Gothic). Allen (1977, pp. 84–85) says that "there are a few examples where the relative pronoun 'attracts' into the case of the head noun phrase", as in (33):

(33)   Ic   wat   wytodlice ðæt ge   secað ðone   haeland
        I   know   truely   that you seek the-ACC savior

        ðone   ðe on rode   ahangen waes.
        that-ACC TE on cross hung   was
                              (Matt 28.5, from Allen 1977, p. 87)

The use of *þe* dies out in the thirteenth century when it is replaced by *þat* (see, e.g., Noack 1992, pp. 21, 35). Forms of *hwa* 'who' start to be used as relative pronouns later in Middle English, at the point when *þat* has taken over from *þe* already. In structural terms, the following change occurs:

(34) a.   CP   b.   CP
          ┌──────┴──────┐   ┌───────┴───────┐
          │   C'   │   C'
          │   ┌──┴──┐   │   ┌──┴──┐
          │   C   . . .   │   C   . . .
       se/þam/   þe   ⇒   who/whom   that
       þat . . .

As I will argue below in great detail, the Specifier that first becomes a head and, after that change, a new *wh*-specifier is introduced for reasons unrelated to Economy but possibly to innovative tendencies (Mustanoja 1960, pp. 192, 194 suggests Latin influence). In effect, these new rein-forcements counter the effects of Economy, as will be seen in the case of negation in the next section as well.

Allen (1977) argues that the demonstrative *se* always takes along the preposition but that *þe* and *þat* never do (see her pp. 83, 76 and 102–105, respectively). This shows that already in Old English, *þat* is in C (her terminology is pre-CP and pre-Spec/head, but I'll put it in current termi-nology), as well as *þe*. This view is similar to Traugott (1972, p. 153), who says *that* the modern relative that is not directly derived from the demonstrative *se*, *seo*, *that*. I will argue that *þat* is sometimes in C and sometimes in the Spec of CP, as a demonstrative, and that the Middle English 'take over' is by the demonstrative *þat* (Mustanoja 1960, p. 188

suggests the same). Showing that *þat* can be a Spec is sufficient to argue against Allen and Traugott, but is not sufficient to argue that the relative complementizer *that* develops from it. I'll give arguments for both.

One of the arguments that *þat* can be in Spec is the presence of combinations such as *þatte*, *þat þe*, and *þat þat*, where *that* is in Spec and the other element in the head position. Grossmann (1906, pp. 26–27) has numerous examples, among them (35) and (36):

(35)   and   wundor   godes   þætte   on   þam   cnihtum
        and   the miracle   of god   that that   to   the   youths

        gecyþed   wæs
        made-known   was

        and God's miracle that was made known to the youths.
                                                           (*Daniel* 470–1)

(36)   eall   þæt þe   styraþ   and   leofaþ
        everything   that   stirs   and   lives
                                                   (Aelfric, *Genesis* 9,3)

That is used not only with neuter antecedents but also with a masculine, feminine, or plural antecedent. This occurs at least from the 9th century on according to Kock (1897, pp. 30–31) who provides numerous examples. This fact can be used to argue either that *that* is already a head and hence not subject to agreement (but that couldn't be the case in (36)) or that *that* even in Specifier position was losing its gender features. This would have made it easier to reanalyze to Head later on.

It is interesting that *that* is followed by *þe* even in non-relatives, as in (37). This occurs frequently, seven times in Beowulf. Zupitza's facsimile edition of *Beowulf* shows *þæt* and *þe* are quite separate even though Klaeber's edition renders them as *þætte* (and even Zupitza's transliteration does so). I take this to indicate that the scribe saw them as two separate words:

(37)   forðam   wearð   ylda   bearnum   undyrne
        therefore became   to-elders to-mankind   not-hidden

        cuð   gyddum   geomore   þæt   þe   Grendel   wan
        known   through-tales   sadly   that   that Grendel   fought

        hwile   wið   Hroþgar
        while   against   Hrothgar

        'Therefore, all mankind found out in sad tidings that Grendel fought against Hrothgar'.   (*Beowulf* 149–51)

This use of *þæt* in non-relatives could be considered evidence that the embedded complementizer *that* also went from Spec to Head. I will not pursue that question here.

Allen's main piece of evidence that *þat* is the C head is that prepositions are never taken along in relatives with *þat* and *þe*, indicating that both move to C not to Spec. There are, however, sentences such as (38), also from the Chronicle and mentioned in Grossman (1906, p. 39) and (39), from Kock (1897, pp. 35–36), that show *þat* is in Spec CP, since *umbe þæt* is a full phrase. This use continues up to the 17th century (see also Seppänen 2000):

(38) þa   þa   hi   þyder comon.   7   umbe oþer þing
      then then they there came.   and about other things

      gesprecon hæfdon.   umbe þ   hi   sprecan woldon
      spoken   had.   about that they speak   wanted

      Then they came there and spoke about other things than which they wanted to speak about.
      (*Chronicle* 1070, Thorpe 1861, p. 344)

(39) þæt is seo lufe embe þæt he wite . . .
      that is the love about that he knows . . .
      (Alfred, *Soliloquiorum* 341: 32 ASC 344 A13)

Bean (1983, p. 92) and Dekeyser (1986) show that the word order with *þe* is mainly SOV. This is explained if the traditional account for Germanic Verb-second holds for OE, namely that the Verb moves to C in clauses where C is not already filled with a complementizer, *þe* in (40). With the demonstrative, it is SVO since the Verb can move to C. Bean (1983, pp. 110–111, n. 3) says that the latter may reflect the paratactic nature of the relative clause introduced by a pronoun. She does not mention *þat* in particular. Due to the rarity of examples with *þat* and considering that V-movement is often optional, it is difficult to obtain evidence from word order for the different status of *þat* and *þe*, but notice that the verb is final in (40), a non-relative, and second in (41):

(40) þa   leton hy   sume. þ þ   mycel unræd wære
      Then some thought that   it would be   great folly

      þ   hy   togedere comon.
      that they should   engage
      (*Chronicle* 1052; Thorpe 1861, p. 314)

(41)   on an igland þæt is ut   on þære sæ   þæt is
on an island that is out on the   sea that is

Meres ig   haten
Mere's island called

(*Chronicle* 895.9, Parker A, Plummer ed. from Bean 1983, p. 91)

Showing that *þat* is sometimes in Spec CP, as in (35) to (41), is not sufficient to argue that the relative complementizer in C in fact derives from the demonstrative in Spec CP. There could be other reasons, e.g., French influence, as Einenkel (1916, p. 119) claims, or analogical extension since Old English *that* is also a complementizer for embedded sentences functioning as arguments or adverbials. An argument for the claim that relative *that* derives from the demonstrative (with neuter singular features) is based on van Gelderen (1997, pp. 76–79).

That is the Old English relative with singular neuter antecedents (Grossmann 1906, p. 38). Hence, the agreement is singular in Old English, as expected. In Middle English, *that* becomes the generalized relative, no longer tied to a singular NP but still with third person singular features: "[t]he Relative (perhaps it does not signify by inflection any agreement in number or person with its antecedent) frequently (1) takes a singular verb, though the antecedent be plural, and (2) the verb is often third person, though the antecedent be in the second or first" (Abbott 1872, p. 167, italics deleted), as in (42) to (48), with the relevant parts in bold:

(42)   and suggeð feole **þinges** . . . þat næuere **nes**
and say   many things   that never   not-was

i-wurðen
happened
(Layamon, *Brut*, Caligula 11472–3, from Mätzner 1864, p. 142)

(43)   **members** that **semeth** lik the maladie of Mirnia
(Chaucer, *The Parson's Tale* 420, from Stoelke 1916, p. 50)

(44)   and it am **I** That **loveth** so hoote Emelye the brighte
(Chaucer, Knight's *Tale* 1736–7)

(45)   **Ye** yeve good counsel, sikirly, That **prechith** me al-day.
(Chaucer, *R. of R.* 5173–4, from Wilson 1906, pp. 47–48)

(46)   we **ladys** and jentil **women** in this contrey that **is** wedows.
(*Paston Letters* III, 338, from Carstensen 1959, p. 83)

(47)   With **sighs** of love that **costs** the fresh blood dear.
(Shakespeare, *Midsummer Night's Dream* III, 2, 97)

(48)   **Thou** that so stoutly **hath** resisted me.
        (Shakespeare, *3 Henry* VI, II, 5, 79, from Visser 1963–1973,
        p. 91)

This agreement pattern can be explained if *that* keeps the features of its
demonstrative ancestor. The 'lack of concord' is due to the features of
the relative complementizer. Even in modern spoken English, 'mistakes'
such as these occur, where *that* rather than the antecedent determines the
agreement:

(49)   There are other **things** you talked about that **is** not on the tape.
        (Christopher Darden, 2 March 1995, "OJ trial")

Sentences (50) to (53), on the other hand, are more modern in that *that*
does not have independent person and number features as a result of its
grammaticalization to head:

(50)   I am the second son of old Sir Rowland, That **bring** these
        tidings.   (Shakespeare, *As You Like It*, V, 4, 159)

(51)   It is **I** That all the abhorred things o'the earth **amend** By being
        worse than they.   (Shakespeare, *Cymbeline*, V, 5, 17)

(52)   But telleth me what myster **men** ye been,
        that **been** so hardy for to tighten heere.
                                              (Chaucer, *Knight's Tale*, 1710–1)

(53)   Two woful wrecches been we, two **caytyves**,
        That **been** encombred of oure owene lyves.
                                         (Idem, 1717–8, cf. Stoelke 1916, pp. 48–50)

Thus, the relative when it functions as subject often triggers the 'wrong
agreement', i.e., singular when the antecedent is plural. Wilson (1906,
p. 45) shows that this does not always take place: "Chaucer's skill in
handling relative clauses referring to personal pronouns of the first and
second persons is shown in that he violates the principle of concord but
rarely". Cases where he does (antecedent and verb are in bold) are (43)
and (44) above. Shakespeare does too, as is shown in (47) above.

The explanation I give is that the relatives start out with features
(singular and neuter), as in (42) to (48), because they are demonstratives
in Spec. They gradually lose these features due to a process of grammat-
icalization, at which point, there is a reanalysis of *that* from Spec CP to
C. The reason for the grammaticalization lies in Economy (as in section
1): it is 'cheaper' to check via head to head movement.

3.1.3. A new Specifier

If Economy is the reason behind the Spec to Head change, the question arises why a new element is introduced in Spec CP again, namely the *wh*-pronoun, as in (54):

(50)　　the est orisonte, which that is clepid comounly the ascendent.
　　　　　　　　　　　　(Chaucer *Astrolabe* Benson, p. 669)

Dekeyser (1986, pp. 100–101) lists the first *wh*-relatives. They appear in the 12th century and are always PPs, as in (55) and (56), quoted in Dekeyser, and in (57) and (58). Allen (1977, pp. 197–199) has similar examples from the Peterborough Chronicle and Homilies, also from the 12th century:

(55)　　hi　　næfdon　na　on　hwam　hi　　fengon　swa　rædlice
　　　　they　not-had　no　on　what　they　caught　so　readily

　　　　'they would have nothing which to seize upon'.
　　　　　　　　(*Peterborough Chronicle* 1085, Clark 1958 edition)

(56)　　þurh　　ungewædera　for　hwan　eor　　þwestmas　wurdon
　　　　through　bad weather　for　why　　their　crops　　　became

　　　　swyþe　amyrde
　　　　very　　damaged　　　　　　　　　　　(Clark 1958, 1110)

(57)　　forðæm　hie　nyton　　mid　hwam　hie　hit　ðe
　　　　because　they　not-know　with　which　they　it　　they

　　　　forgielden
　　　　recompense
　　　　　　　(Alfred, *Pastoral Care*, Sweet 1871, hence PC, 323.23–4)

(58)　　ðonne　mon　mæg　ongietan　of　hwam　　hit　æresð　com
　　　　then　　man　may　understand　whence　it　first　came

　　　　& and　for　hwæm.
　　　　and　　why　　　　　　　　　　　　　　　(Idem 241.16)

In Old English, the preposition is never stranded (e.g., Allen 1977), and in cases where it looks like stranding, as in (59), these involve a head complementizer, i.e., not one moved as the object of the preposition:

(59)　　Seo　gesyð　ðe　we　　god　mid　geseon
　　　　the　sight　that　we　see　God　with
　　　　　　　　　　　　　(Alfred, Sol. 67.6 from Allen 1977)

The reason behind introducing *wh*-relatives in constructions such as (55)

to (58), rather than as subject or object relatives, can be explained in a functional manner: in order to relativize a PP, a Specifier position is needed (e.g., *hwam* in (58) is part of a full phrase and does not move on its own). Thus, *wh*-elements fill the gap that is left by *that* becoming a Head.

As to what the origin of the new specifier is, as mentioned, Mustanoja cites Latin influence, and Rydén (1983) shows both Latin and French influence. Rydén shows that the first instances of *who* occur in epistolary idioms that are very similar to those in French letters of the same period. For instance, in many of the collections of letters from the fifteenth century, the same English and French formulaic constructions occur, such as in (60) from the writings of Bekynton and (61) from those of Paston:

(60)   a   laide   de Dieu notre Seigneur, Qui vous douit
        with the-help of God our   lord,   who us   gives

        bonne vie et   longue.
        good   life and long   (Bekynton, from Rydén, p. 131)

(61)   be the grace of God, who haue yow in kepyng
                                        (Paston #410, Davis, p. 655)

Bergs (2002), based on Rydén, argues that *who* is initially restricted to a deity antecedent (as in (61)), constituting a lexical innovation. This fits with *wh*-relatives being introduced for reasons of expressivity, i.e., for language external reasons.

In conclusion to section 3.1, the relative head *that* develops out of a specifier in accordance with (4), providing a structural explanation for this kind of grammaticalization. Other specifiers (*who/m*, *which*) are introduced for reasons of expressivity. Thus, Economy interacts with innovations, making cyclic change possible.

3.2. *Negation*

In this section, I describe a well-known change that shows many of the same stages as the relative pronoun change just examined. I start by providing some theoretical background on the NegP, giving a description of Jespersen's Cycle, in terms of the Heads-over-Phrases Economy Principle (4), and indicating briefly how prescriptive rules stop it from further change in Modern (standard) English.

Multiple negation is found in many languages, from Standard French, Navajo, Yiddish, and Celtic to Afrikaans and Mayan (Dahl 1979, p. 88). This phenomenon is explained elegantly if negation is located in a NegP (as in Zanuttini 1991 and Ouhalla 1990; or a ΣP as in Laka 1994), with

a Specifier that checks its negative features with a Head (also known as the Neg-Criterion, similar to the *Wh*-Criterion discussed above). In many of these languages, one of the negative parts is attached to the finite verb, indicating it is a head. Negative determiners in these languages often trigger the presence of the negative marker (i.e., they are negative polarity items, hence NPI). In languages such as Italian and African American English, the doubling takes place only when the N-word is in the scope of the negation.

Jespersen (1917) describes what has come to be known as Jespersen's Cycle, where historically the negative element attached to the verb disappears, and the other negative element takes over. Latin *non dico* goes to *jeo nedi* to *je ne dis pas* and *je dis pas*. These are cases of specifiers becoming heads. Using (4), Jespersen's Cycle can be accounted for straightforwardly. Assume negation is checked in a NegP, as in (62). The change then involves grammaticalization whereby a head *ne* becomes a clitic *n-* (prefixed to a verb). This change is followed by the introduction of a new XP, such as *no thing*, in specifier position:

(62)   NegP

                             Neg′

                      Neg

        no thing   [+NEG]

                  ne
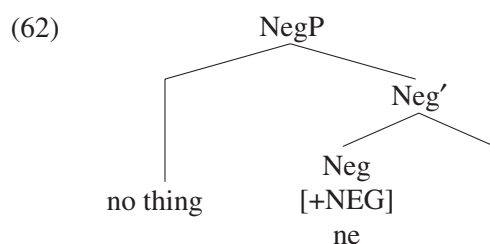
Subsequently, the XP in the specifier becomes a head *not,* and the above changes could start over again. What stops the cycle is the prescriptive rule (perhaps the most prevalent of all) against multiple negation since for the specifier to be filled again, there would have to be a second negative. For some of the admonitions against the use of multiple negation, see Kirszner and Mandell (1992, p. A37); Quirk and Greenbaum (1973, p. 186); Swan (1980, p. 182). These rules have existed at least since the 18th century and are very strong. Nevertheless, speakers continue to produce them. For instance, Anderwald (2002, pp. 104 ff) examines the BNC and finds a considerable number of multiple negation in spoken British English (namely 14.3% of negatives).

In recent literature (e.g., the articles by Rissanen and van Kemenade in the 1999 volume on *Negation in the History of English* edited by Tieken-Boon), the emphasis has been on verb-movement and negation, the decline of multiple negation, etc., but not on the phrasal origin of *not*. For instance, van Kemenade (1999, p. 152) examines *na/ni* and not variants of *nawiht*

'no creature' because the latter "often function as modifiers, but their use as such is often difficult to separate from their use as a negated noun . . . I leave more detailed consideration of these for further research". Hence, she does not consider cases where the negative functions as a determiner, which I do in sentences such as (65) below. In what follows, I describe the English situation where *nan wuht* 'no creature/person' takes over the function of the head *ne*, but first I'll describe the situation in contemporary Germanic.

The Germanic negative *not/nicht/niet* develops out of a form of *na wight/ne ie wicht* 'no creature'. However, linguists are not in agreement as to whether the modern forms are heads or specifiers. For instance, Haegeman (1995, pp. 119, 191) suggests *nicht* in German is a specifier, against Bayer (1990), who assumes it is a head. Wood (1997) finds it very difficult to find evidence in 15th century English. Afrikaans *nie* and Dutch *niet* are also assumed to be specifiers by Haegeman. Yet, in the majority of instances, these negative elements (seem to) adjoin to the verb (the first one of those in final position), as in (63), from Dutch:

(63)   'K   heb   'm   gisteren   nie(t)   gezien
        I   have   him   yesterday   not   seen

Haegeman shows that there are cases where *nie* is a specifier, e.g., (64) from Flemish, and then argues that if it is a specifier once, why not all the time:

(64)   da   Valere woarschijnlijk nie   styf   drunke (en)-was
        that Valere probably   not   very   drunk   NEG-was
                                                (taken from Haegeman 1995)

In Flemish, there is negative concord with an optional *en*, as indicated in (64). *En* cliticizes to a verb (and functions like *ne* in French). Middle Dutch has a similar negation (and so does Old English, as will be shown below). It is therefore expected that *en* in (64) is the head and *nie* the specifier. In Dutch, Standard English (see section 4.1), and Standard German, where negative concord does not occur, it seems to me that the negative element is ambiguous between Spec and Head. I will now focus on the history of English and will first discuss the variant forms of negation and the status of the negated elements. Then I will move to the changes from Specifier to Head and see how they interact with non-linguistic factors.

In Old English, as mentioned briefly in (62), the regular negation is *ne*, which is often contracted with the verb, as in (65), especially in southern texts such as the Pastoral Care (abbreviated as PC):

(65) Ac   nis   nan   scild   trum[ra] wið   ðæt tuiefalde
      but not-is   no   shield stronger against the twofold

      gesuinc ðonne mon sie untwiefeald
      toil   than   man be   sincere

                                     (Alfred, PC 239.9 Hatton)

I will assume this *n-/ne* is a head and that the finite verb moves to it on its way to C, Old English being Verb-second, as in (66), where I ignore most details:[12]

(66)
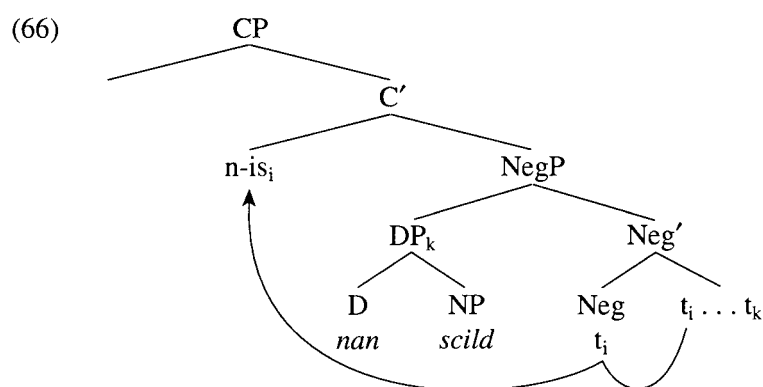
Once *ne* weakens phonologically, another negation is introduced, one that grammaticalizes from a full XP, as in (67) and (68), to a head. Many variants of these forms coexist, e.g., *nauht* and *nan wuht*, as well as *na* and *no*, and *ænig wiht/litel wiht*, meaning 'a little' (comparable to the French *pas* 'step'):

(67) forþæmpe hie   hiora nan wuht   ongietan   ne   meahton
      because   they their no   thing understand not could

      because they couldn't understand anything

                                     (Alfred, PC 4/12 Cotton)

---

[12] Frisch (1997, pp. 30 ff.) argues that *not* in Old English and Early Middle English is not in NegP, but is a sentential adverb like *never*. It seems to me that, if correct, this should result in negatives cancelling each other out. They don't and that's why a NegP makes sense.

(68)   ðylæs hie   aught   sellan ðæm þe   hie   nanwuht ne
        lest they anything give   them that they nothing   not

        scoldon
        should

        lest they give anything to whom they should give nothing
                                                    (PC Cotton 320/15–6)

The next step in the grammaticalization process is for the negative NP
to become a negative adverbial, as in (69), and to become one word
phonologically:

(69)   Næron   ȝe   **noht** æmettiȝe,   ðeah   ge   wel   ne
        not-were you not   unoccupied. though you well not

        dyden
        did

        You were not unoccupied, though you did not do well
                                                    (PC Cotton from the OED)

The element moving to Spec can be a full phrase. In fact, both *nane wuht*
and the more grammaticalized *nawht*, or variants thereof, as in (69) and
(70), function both as negative adverb and as argument:[13]

(70)   Ac   nis   nan   scild trum[ra] wið   ðæt tuiefalde
        But not-is no   shield stronger against the   twofold

        gesuinc ðonne mon sie   untwiefeald, forðremðe nawuht
        toil   than   man be   sincere,   for   nothing

        nis   ieðre to [ge]secganne, ne eac to [ge]liefanne
        not-is easier to say   and   believe

        ðonne soð
        than   truth   (PC, 239.9–10)

   If the Modern English adverb head *not* arises out of an Old English
noun phrase, when does this first start happening? The Helsinki Corpus
(abbreviated HC) gives a good indication as to first use (even though it

---

[13]   There are cases where the negative is in subject position in OE, and eME:

   (i)   forðremþe nan wuht nis betere
          because   nothing is   better   (Alfred, *Boethius* 85/1)

   (ii)   þat nan   ne   beo so wilde
          that none not is   so wild   (Layamon, *Brut*, Caligula 395)

I'm not really sure of the correct analysis for these.

cannot be conclusive). The data are provided in Appendix 1. In OE 1–2 of the HC (see Appendix 1 for periods of the HC), there are no instances of *nanwuht* used in the modern sense, i.e., as a negative adverb. All seem to be either DP or D.[14] In OE3, there is one possible instance of modern use, namely (71). The more likely explanation, however, is that it is an argument DP, not an adverbial one. In OE4, there is one instance, namely (72), but again there is a reading with *naht* as argument. The OED lists some early ones, as in (69) repeated here as (73):

(71) Ne biþ   us   frea milde . . . gif we yfles   noht
      not is [he] to-us lord mild . . . if   we of evil nothing

      gedon habbaþ
      done   have   (*Juliana* 328–9)

(72) Nis   þin   mæʒn nabt wið   hire forþon þe . . .
      not-is your power not   against her   because . . .
                                          (*Passion of St. Margaret* 172/75)

(73) Næron   ʒe   noht æmettiʒe,   ðeah   ge   wel ne
      not-were you not   unoccupied though   you not did

      dyden
      well

      You were not unoccupied, though you did not do well.

                                                     (PC Cotton)

From ME1 on, as shown in (74) and (75), the NP is used as negative adverb. Some of these are still regular arguments and occur with *ne*, as in (76). Others, such as (74) and (75), are non-argumental and are like modern *not*. The OED says that *nought* and others are used in the meaning 'not at all'. In ME2–4, *nawuht* is replaced by a monosyllabic form:

(74) ne ne   helpeð nawiht eche   lif   to haben
      doesn't help   not   eternal life to have
                                                   (Katherine (1230) 26/6)

(75) for   he ne mei nawiht luuie god; and ec   his
      because he not can not   love God and also his

      ehte
      wealth   (Lambeth Hom, 147/11)

---

[14] There is, however, an emphatic negative adverb *no/na* used (with optional *ne*) from early DE on. This occurs 43 times in Beowulf. This adverb could have been the one to take over when *ne* disappeared but didn't. I won't go into this further (see e.g. van Kemenade 1999, 2002).

(76)   ne   drede we nawiht
        not dread we nothing   (Sawles Warde, 175/10)

Negative NPs, as in (76), and negative adverbs, as in (75), are often
identical in form and, assuming the NegP, they would be situated in the
same (Specifier) position, as in (88) above. Hence, the NP comes to be
used as a general negation.

The forms *nat/not* are first used in Late Middle English: *not* is first used
in 1362, according to the OED (see also Appendix 1), still with *ne*. I will
not go into that. As to when *not* starts to lose its phonological indepen-
dence, according to the HC, Farquhar and Vanbrugh (both writing
comedies in the late 17th century) are the first to use forms such as *won't*
and *don't*, and they are the only authors in EMod 3 with a total of 56 forms,
as in (77). None of these is followed by another negative, however:

(77)   I won't dispute it now. (Farquhar, The Beaux Strategem (1707),
        p. 7)

(78)   Don't you be jealous now. (Vanbrugh, The Relapse (1696),
        I 34)

Jespersen (1917, p. 117) argues that forms such as *donot* start around 1600.
Rissanen (1999, p. 196) has evidence from 1500 on of sequences of
[auxiliary + *not* + pronoun] in questions, which indicates a close rela-
tionship between the auxiliary and the negation. It is not clear when this
reduced form of *not* is first reinforced in (non-standard) English by another
XP such as *nothing*. Cheshire (1999, pp. 30 ff) suggests that early on *never*
was ready to take on the task of reinforcing the negative and would have
done so "[i]f the process of standardisation had not intervened" (p. 31).
Jespersen (1917) says that multiple negatives become used to portray
lower-class by the 17th century. So, the prescriptive prohibition against
multiple or double negation starts early and, as mentioned above, is
perhaps the strongest of all prescriptive rules. This is the reason a new
Spec is not introduced in standard English and the cycle is interrupted,
and the head *n't* is retained as well.

In summary, the changes that take place in negation can be seen as (a)
to (d). Important for the present paper is that XP becomes X and that the
XP is not just any XP but an N-word that needs to check its features with
a Negative Head:

(a) Early Old English has *ne* in head position, with *nan wuht* optionally
     present in Spec of NegP; this two-word form dies out by Late Middle
     English (see Appendix 1);

(b) Reduced forms such as the two-syllable *nawuht* and the monosyllabic *naght* occur in Old English. In ME1, the bisyllabic form still occurs but ceases to do so in ME 2–4. Both it and the monosyllabic form occur accompanied by *ne*.

(c) *nat/not* are first used in Late Middle English (see also Appendix 1), still with *ne*;

(d) *not* is used by itself by the early Modern period and is then a Head that starts to contract to *n't* in C17/C18. The expected new negative reinforcement in Spec becomes stigmatized in C18. It does not occur in standard English, due to tremendous prescriptive pressures, but is pervasive in vernaculars.

There are more specifier candidates in Old English, and the question arises why *no wuht* is 'chosen'. For instance, *nan wuht* is in competition with *nan þing*. The reason the former 'wins out' is pure chance (see van der Wouden 1994 about collocational behavior). It happens to be phonologically reduced earlier than *nan þing*. There are two indications for this: (a) looking at the collocates of *þing* shows more variety than with *wuht*, and (b) while forms such as *naþing* do not start to occur until ME1 (OED says 1200 for *naþing*, 1390 for *nothing*), noncontracted forms such as *nan þing* continue to occur much later:

(79) þat no bið he for   þan watere. nað|ðing[15]
      so *not* *be* *he* *because* *the* *water* *nothing*

      idracched
      damaged

      So he is not at all damaged by the water (Layamon, Caligula 11002, from OED which says it means 'not at all')

In tree form, the changes discussed above as (a) to (d) are shown in (80a–d), with the change from (c) to (d) most relevant to the XP to X change:

---

[15] The '|' indicates that the line ends there and the consonant is doubled because of it.

(80) a.

```
                    CP
           ┌────────┴────────┐
                            C′
                   ┌─────────┴─────────┐
                 n-isᵢ                NegP
                          ┌────────────┴────────────┐
                        DPⱼ                         Neg′
                   ┌─────┴─────┐              ┌──────┴──────┐
                   D          NP            Neg          tᵢ . . . tⱼ
                   na        wuht            tᵢ
```

b.

```
                    CP
           ┌────────┴────────┐
                            C′
                   ┌─────────┴─────────┐
                 n-isᵢ                NegP
                          ┌────────────┴────────────┐
                        DPⱼ                         Neg′
                         D                   ┌──────┴──────┐
                      na(w)uht              Neg          tᵢ . . . tⱼ
                                             tᵢ
```

c.

```
                    CP
           ┌────────┴────────┐
                            C′
                   ┌─────────┴─────────┐
                 n-isᵢ                NegP
                          ┌────────────┴────────────┐
                        APⱼ                         Neg′
                         A                   ┌──────┴──────┐
                       nat/not              Neg          tᵢ . . .
                                             tᵢ
```

d.

```
                    CP
           ┌────────┴────────┐
                            C′
                   ┌─────────┴─────────┐
                                      NegP
                          ┌────────────┴────────────┐
                         ∅                          Neg′
                                             ┌──────┴──────┐
                                            Neg           . . .
                                          not/n't
```
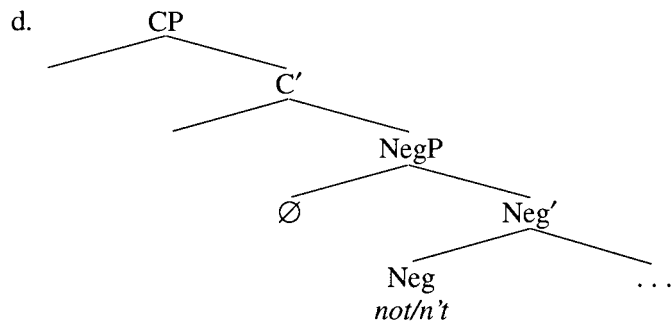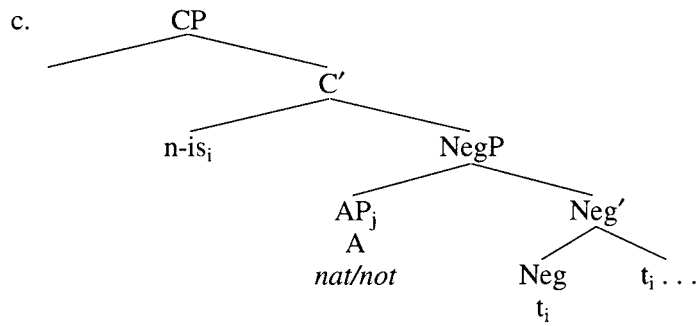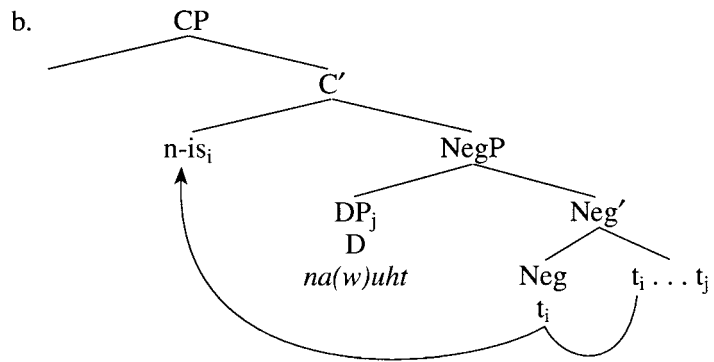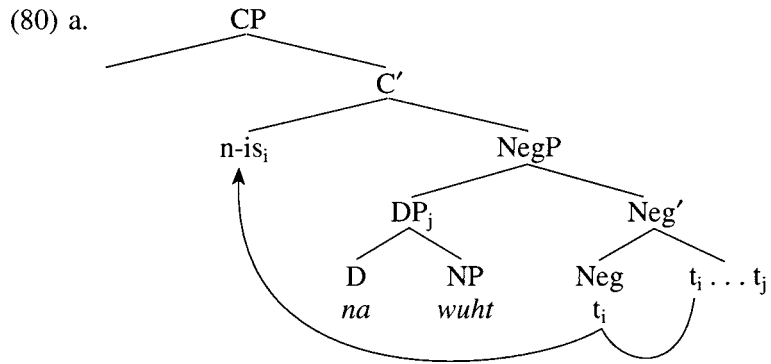
3.3. Whether

Whether has a number of functions throughout the history of English. In some uses, it starts out as a specifier and either disappears or is 'forced' (by prescriptive rules) to remain a specifier. The discussion again illustrates how the general principle in (4) interacts in a variety of ways with other changes in the language.

   Whether introduces a direct question, as in (81) to (84). In this yes-no question function, *whether* causes the verb to be in second position, even as late as the 18th century. This shows *whether* is in the specifier position of CP and the verb in C:

(81)   Hwæðer wæs Iohannes   fulluht þe   of heofonum þe
       Was   the   baptism of John   that of heaven   or

       of man.
       of man   (*AS Gospel* Matthew 21.25)

(82)   Hwæðer wille ge   ðæt ic cume to eow, ðe mid
       Do   you   want that I   come to you,   with

       gierde ðe mid monðwære gæste?
       a rod   or with gentleness   of spirit?
                                        (*Alfred Pastoral Care* 117.7–8)

(83)   Whether hadst thou rather be a Faulconbridge, . . .
                                        (Shakespeare, *John* I, i, 134)

(84)   Whether doth doubting consist in embracing the affirmative or
       negative side of a question?
                                        (Berkeley, *Hylas* I, 173, 10 from 1713)

After the 18th century, *whether* disappears, and verb movement to C suffices. Assuming question features are checked in CP (also referred to as the *Wh*-Criterion, see Rizzi 1990, pp. 65ff), this is initially done via Spec-head agreement but later by Head-checking of the verb. So, this is a case where (4) has to be seen as a condition on checking.

   Whether has a number of other uses, most prominently as pronoun and as complementizer. (85) and (86), taken from the OED, show pronominal use. Like the direct question *whether*s the pronouns are always in the specifier position in OE and ME since they are phrasal and trigger Verb-second:

(85) Hwæðer þara twezra dyde þæs fæder willan?
Which of the two did the father's will?
(*Ags. Gosp. Matthew* xxi. 31)

(86) hweðeres fere wult tu beon?
Whose friend do you want to be?
(*Ancrene Riwle* 284)

This pronominal use of *whether* disappears gradually and is replaced by *which*.

Complementizer *whether*, as in (87) to (89), has survived up to now and is initially either the Specifier (in (87) and (90) to (92), as shown by the verb in C) or in the Head C (in (88) and probably in (89)):

(87) nast ðu hwæðer beoð þæs rican mannes ban.
not-know you whether be that rich man's bone,

Hwæðer þæs pearfan
or that poor (*Aelfric Homilies* I Thorpe 256)

(88) ða cwædon . . . hwæðer ænig man him mete
then said . . . whether any person him food

brohte
brought (*AS Gospel* John 4 33, DOE, seg 25)

(89) I know not whether Heauen will haue it so. (Shakespeare *1H4*)

(90) whether that alle thise thinges maken . . . (HC ME3)

(91) whether that I lyve (HC ME4)

(92) Then iudge . . . if I haue done amisse: Or whether that such Cowards ought to weare This Ornament of Knighthood, yea or no? (Shakespeare, *1H6*)

Under (4), one would expect *whether* to become C in present day English. In the case of heads, e.g., *that* in (93), *wh*-movement moves via the embedded Spec CP, and the sentence is grammatical. This is not so with *whether*; it isn't considered a head since *wh*-movement across it is blocked, as (94) shows:

(93) Who did you think that he met?

(94) *Who did you wonder whether he met?

Snyder (2000) reports on a psycholinguistic experiment in which speakers who hear (94) start accepting the construction as grammatical, meaning that they could analyze *whether* as a head with the right trigger.

The reason, however, that *whether* has not become a head, I will argue, is prescriptive pressure. For instance, Kirszner amd Mandell (1992), in their writing guide, say that *whether or not* is used 'when expressing alternatives' (1992: A26) and is then forced to be a phrase and a Spec.[16] In the CSE, *whether* is immediately followed by *or not* in 18% of the instances (to be precise, there are 257 instances of *whether or not* and 1,141 of just whether). There is also direct evidence to the speaker that *whether* is in Spec since sequences such as (95) and (96) are encountered, even in formal speech, as in (95), and writing in (96):

(95)   I just wondered whether that as a next step we might look to see why this seems to be the case (CSE-FACMT97)

(96)   The local authority will know whether if they let the council house to the tenant. (BNC-FC3-80)

Regarding *whether*, historical changes show interesting variety. In the case of yes-no questions, the specifier *whether* disappears, and the verbal head suffices to check the features, in accordance with (4); in the case of embedded yes-no questions, *whether* has remained a specifier. I argue that the latter is caused by prescriptive pressures. If *whether* needs to be followed by *or not*, it has to be a specifier.

In summary to section 3, I have provided evidence that three changes in English can be described as changes from phrase, i.e., spec(ifier), to head (XP > X). The reason for that change is (4) above, an Economy Condition that says 'project as a head if you can'. Counteracting these changes by reintroducing new specifiers are innovations such as *wh*-relatives and negative reinforcements. Factors holding back the development from Spec to Head are prescriptive rules such as 'do not use a double negative' and 'never end your sentence with a preposition'. I will now examine changes involving lexical heads that also make use of an Economy Principle, but a different one.

4. 'Merge Late'

In work on grammaticalization and reanalysis, there has been a lot of emphasis on the change from lexical to auxiliary verbs (e.g., *have* and *will*), from prepositions to complementizers (e.g., *for* and *like*), and from

---

[16]   Another reason could be 'analogy' with the other *wh*-complementizers, such as in *I don't know what they saw*.

verbs to complementizers. As a lexical head is reanalyzed as a grammat-
ical (or functional) head, it 'climbs' up the tree, as it were (see also IJbema
2002). Modals (changing from V to Auxiliary), perfect marking (from V
to Auxiliary), progressive (from P to I), infinitive markers, and comple-
mentizers (van Gelderen 1998) all fit into this.

Chomsky's (1995, 2001a, 2001b) preference of 'Merge over Move',
reformulatable as 'Merge late so that you don't have to merge as well as
move', as in (97), repeated from (6) above, provides a ready explanation:

(97)   Late Merge
        Merge as late as possible.

Roberts and Roussou (1999) use it to explain the change from V to AUX.
At the time of the change of main verbs to auxiliaries, verbs move to I
and/or C to check features of tense and agreement, and having the auxil-
iary merge rather than merge and move in e.g., (98) is more economical:

(98)   ne   **mahte** he wið leasse gref **habben arud** us?
        Not could   he with less   pain have   saved   us?
                                                (*Ancrene Wisse*, 106r)

I will use this principle of Late Merge to account for some of the other
Head-to-Head changes. In 4.1, I'll explain the change from P to C using
(6) and show how the contemporary change of *to* from I to C is stopped
both by (4) and by a prescriptive rule. Condition (6) also explains the
change from a lower Spec to a higher Spec, as I'll show in 4.2.

4.1. Head to head

As is well-known, the preposition of location and purpose, as in OE (99)
to (101), comes to be used as a complementizer by early ME, as in (102)
and (103):

(99)   þæt he **for**   eaxlum ʒestod
        that he before shoulder stepped   (*Beowulf*, 358)

(100)   **for** werefyhtum . . . ond **for** arstafum   usic sohtest
        for fighting . . .   and for support (you) us   sought
                                                (*Beowulf*, 457–8)

(101)   **forþan**   ic hine sweorde swebban nelle
        therefore I   him   sword   kill   not want
                                                (*Beowulf*, 679)

(102) Locrin 7   Camber to þon scipen comen. **for** to
       Locrin and Camber to the ship   came   for to

       habben al þa æhte
       have   all the property   (*Brut*, Caligula 1113–4)

(103) moche he lofde echn(e) cniht. þat lofde **for** to
       Much he loved every knight who loved to say

       segg(e) riht
       the   truth.   (*Brut*, Otho, 5523)

On first sight, these don't seem to be cases where the P moves to C, as
with verbs such as in (98) above, and where Merge would be simpler than
Move. The explanation I give follows Kayne's (1999, 2002) account of
certain prepositions. He argues that prepositions such as 'of' are merged
outside the VP. Adapting that to (99)–(103), I argue that once a preposi-
tion is no longer fully lexical but is used to express (purposive) Case, as
in (100), it is merged outside the VP (with the NP inside the VP) and serves
as an attractor to its NP 'object'. Once the P is base generated outside
VP, it can be reanalyzed as a complementizer indicating purpose, e.g., in
(103). This accounts for the change from P to C.

There are some problems with 'Merge over Move'. For instance, why
couldn't one merge all the time, i.e., always have *there* subjects and *do* in
I? In the remainder of the paper, I provide some evidence that *to* moves
from I to C, but 'prefers' to stay merged in I, not C. The reason for this
'early Merge' is a conflict with (4).

The ME changes involving *to* from P to I have been discussed many
times, and here I will examine some idiosyncracies of *to* and *not* in this
light (I won't go into its precise position, but see van Gelderen (2002)
and Abraham (2002)): As is well-known, the split infinitive has occurred
since the 14th century, but examples with a negative, as in (104), are less
frequent than when a non-negative adverb is involved and (105) is pre-
ferred:

(104)   He professes to not be ready for that (BNC-CGB 1649)

(105)   He professes not to want the job (BNC-ABJ 970)

For instance, in the CSE, there are 381 instances of *not* preceding infini-
tival *to* and 59 of *to not*, indicating that only 13.4% of negative infini-
tives are split whereas with IP-adverbs *probably*, *perhaps*, *possibly*, this
is 71% (see van Gelderen 2002). I will show that there is a position for
to above the negation (C) and one below it (I). Since the one below it is

selected, 'early Merge' seems preferred over 'late Merge', but only with negative adverbs.

One of the arguments for two positions is that in a number of modern corpora (CSE, BNC, Cobuild), (106) and (107) occur, and when asked, native speakers say they've heard them and are somewhat acceptable:

(106)   – as a request to not to –. (CSE-WH97A)

(107)   This is to try to not to overturn the . . . (CSE-WH97A)

This shows there are two positions for *to* to occur in: a lower one as in (105) and a higher one as in (104). Since the preferred position of *to* is after *not*, there are several possible accounts, but I argue that *to* moves to C in (104) leaves a copy in (106) and (107).

Sentences with an overt complementizer, such as (108), occur and provide evidence that *to* is in C, in complementary distribution with *for* (for instance, in the BNC, there is one *for N to not* in a very muddled sentence but 126 *for N not to*):

(108)   if you as a parent choose for your child not to participate, you may opt out (CSE-COMR6B97)
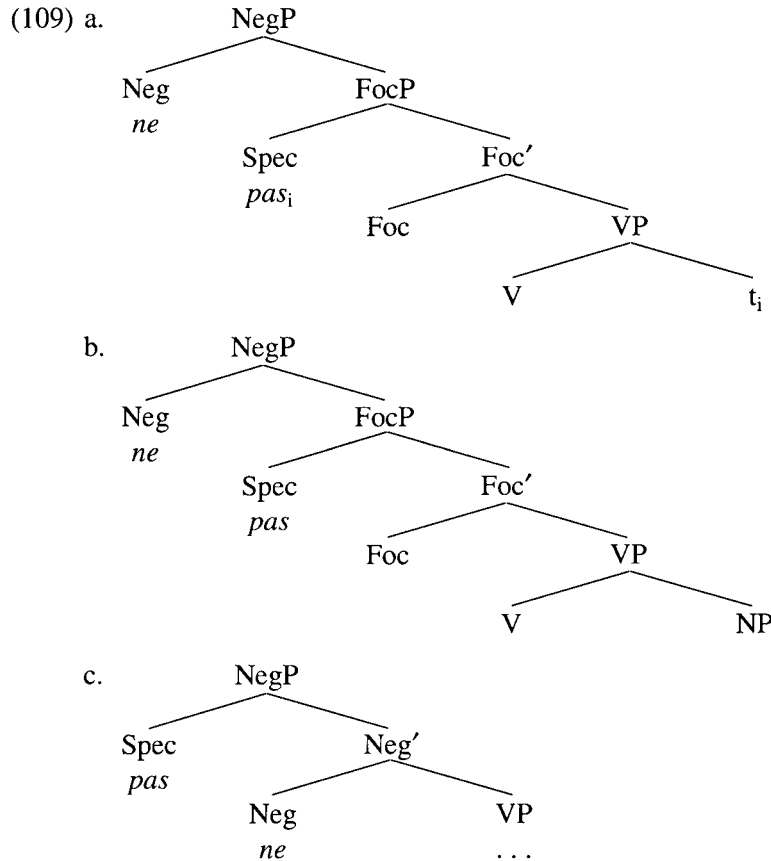
Verbs without a CP complement, such as *believe*, are judged grammatical by native speakers with *not to* but not with *to not*, further indicating that *to* in the latter combination is in C.

The problem for the 'Merge over Move' principle is to explain why *to* would not always be in the higher position. There are two possible reasons. One is that the rule against splitting infinitives is still very strong in standard English (see Fowler 1926 [1950], p. 558; Quirk and Greenbaum 1973, p. 312). The second is that typically *not* is a head (see 3.2 above) and that in order to move across *to*, as in (106) and (107), *not* is forced to be a specifier. This is still possible in Modern English since nothing else is in the Specifier position (due to the double negation prohibition). However, that makes the move infrequent enough not to trigger late Merge. So, two Economy Principles interact, and one 'bleeds' the other.

4.2. Spec to Spec

A change that seems to occur less frequently than Head to Head or Spec to Head but which is accounted for by (6) as well is the change from specifier of a lexical category such as VP to a functional one. Simpson and Wu (2002, pp. 291 ff.) analyze negation in the history of French in terms of negative *ne* selecting a Focus projection below NegP but above

the VP to which the negative object moves, as in (109a). This object then becomes base generated in the Spec of FocP, as in (109b), and subsequently in Spec NegP, as in (109c):

(109) a.

```
              NegP
            /      \
        Neg         FocP
        ne        /      \
              Spec         Foc′
              pasᵢ       /      \
                      Foc        VP
                              /      \
                            V          tᵢ
```

b.

```
              NegP
            /      \
        Neg         FocP
        ne        /      \
              Spec         Foc′
              pas        /      \
                      Foc        VP
                              /      \
                            V          NP
```

c.

```
              NegP
            /      \
        Spec        Neg′
        pas       /      \
              Neg          VP
              ne           . . .
```

Spec to Spec falls under the 'Merge over Move' principle since in (109b) there is less movement than in (109a), and the negative is merged latest in (109c).[17] (As in English, the next step will be for *pas* to become a head, in accordance with (4). This has presumably happened in varieties of French where *ne* has disappeared.)

In section 3.1 above, I examine relatives and argue that phrases in the specifier change into heads. The sequence of events for relatives is: (a) *that* is in Spec and *þe* in the head of CP, as in OE (32) above, (b) *that*

---

[11] Roberts and Roussou (1999) examine this change, skipping the focus stage, but use the Lexical Subset Principle (if an element always occurs in one environment: reanalyze it) to account for the change to (109c).

becomes the head, and (c) the specifier of CP is filled again by a *wh*-element that moves there, as in (54) above. There is, however, a short period from 1400 to 1600 where *wh*-pronouns are base generated (i.e., merged) in Spec of CP, as in (110) and (111), with a resumptive pronoun in argument position:

(110)   þis is he, which þat myn vncle swereth he mot be ded
          This is he, who my uncle swears must be dead.
                                    (Chaucer, Troilus II 654, from Visser p. 58).

(111)   I tell my sorrowe to the Stones, who though they cannot answer
          my distresse Yet in some sort they are better than the Tribunes.
                                    (Shakespeare, Titus, III, i, 37, also from Visser)

Constructions (110) and (111) represent stage (d): the *wh*-element is base generated (merged) in the Spec of CP, and this is expected given (6). If a *wh*-element merges inside the VP and then moves to Spec CP, it is more economical to merge it late. The reason, I argue, that stage (d) is short and that constructions such as (110) disappear is that they interact with (4), the change towards heads.

In section 4, I have provided a few instances of changes that follow the late merge condition as formulated in (6). Head to head changes discussed in this context are verbs becoming auxiliaries and prepositions becoming complementizers. The change of *to* from P to C is hindered, however, by the ban against split infinitives, as well as the interaction with Condition (4). Two changes of specifiers becoming specifiers (of higher functional projections) follow the same condition to undergo merge as late as possible. In the case of the relative clause change, there is again an interesting interaction with principle (4) and the latter 'wins out'.

5. Conclusions

The main part of the paper (section 3) is concerned with instances of phrases that become heads. Economy of a head over a phrase, principle (4), is seen as the motivating factor for changes involving *wh*-constructions and negatives and nicely describes pronominal behavior (section 2). Figure 1 shows that Spec-Head checking is replaced by Head-Head checking, with the Spec becoming the head and the original head deleting. Subsequently, a new Spec is introduced which in turn will become a head. The introduction of a new Spec is not motivated by Economy factors:
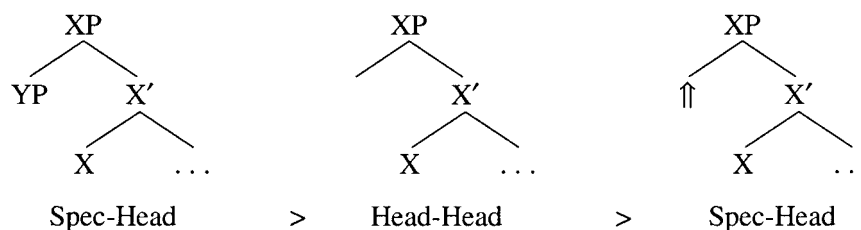
Figure 1. Spec to Head Cycle.

In all of the changes described,[18] the Economy Principle interacts with introductions into the language of new items, e.g., *who* and *na wiht*. The latter occur to increase expressivity and are the result of 'chance', of language-external changes. The changes involving Spec to Head are changes caused by language-internal factors.

Instances of head to head grammaticalization can be explained by late merge: if a head is grammatical rather than lexical, it can merge outside the VP. In cases where late merge does not apply, e.g., in *to* changing from I to C, there are other factors at work. Changes from Spec to Spec occur as well but are perhaps harder to document. Possible cases are French negatives and a stage in English relative clauses.

Appendix 1

	'*nun wuht*'	'*nawuht*'	'*naght*'	*not/nat*[19]
OE I-II (–950)	10	24	35	0
OE III (950–1050)	1	5	42	0
OE IV (1050–1150)	0	1	23	0
ME I (1150–1250)	3	53	313	1/10
ME II (1250–1350)	1	0	144	5/0
ME III (1350–1420)	9	0	220	695/282
ME IV (1420–1500)	1	0	379	788/83

Variant forms of nan wuht, nawuht, naught, and not in the Helsinki Corpus.

---

[18] Spec to Head changes I have not examined are preposition-stranding in general (not just in *wh*-constructions) and the 'article-cycle', as in Lyons (1999). There is interesting evidence in German, for instance, that the original specifier *Anna* in *Annas Buch* is now a head (see Rosenbach 2002, p. 209 and Demske 2001).
[19] Note that *nat* occurs often in OE but with the meaning 'know not'. I have obviously not included these.

References

Abbott, E. A.: 1872, *A Shakespearean Grammar*, MacMillan and Co, London.

Abraham, Werner: 1992, 'Vorwort; Einleitung zum Bandthema', *Folia Linguistica Historica* **13**(1–2), 7–26.

Abraham, Werner: 2002, 'The Grammaticalization of the Infinitival Preposition: Towards a Theory of "Grammaticalizing Reanalysis"', ms.

Allen, Cynthia: 1977, *Topics in Diachronic English Syntax*, unpublished Ph.D. dissertation, University of Massachusetts –Amherst.

Anderwald, Lieselotte: 2002, *Negation in Non-Standard British English*, Routledge, London.

Bayer, Joseph: 1990, 'What Bavarian Negative Concord Reveals about the Syntactic Structure of German', in J. Mascaró and M. Nespor (eds.), *Grammar in Progress*, Foris, Dordrecht, pp. 13–23.

Bean, Marian: 1983, *The Development of Word Order Patterns in Old English*, Croom Helm, London.

Benson, Larry: 1987, *The Riverside Chaucer*, Houghton Mifflin Company, Boston.

Benveniste, Emile: 1966, *Problèmes de linguistique générale*, Gallimard, Paris.

Bergs, Alexander: 2002, 'Grammaticalizing Relative *who*', talk delivered at New Reflections on Grammaticalization meeting, Amsterdam.

Cardinaletti, Anna and Michael Starke: 1996, 'Deficient Pronouns: A View from Germanic', in H. Thráinsson et al. (eds.), *Studies in Comparative Syntax II*, Kluwer, Dordrecht, pp. 21–65.

Carstensen, B.: 1959, *Studien zur Syntax des Nomens, Pronomens und der Negation in den Paston Letters*, Bochum-Langendreer.

Chapman, Carol: 1995, 'A Subject-Verb Agreement Hierarchy', *Historical Linguistics* **2**, 35–44.

Cheshire, Jenny: 1999, 'English Negation from an Interactional Perspective', in I. Tieken-Boon et al. (eds.), *Negation in the History of English*, Mouton de Gruyter, Berlin, pp. 29–53.

Chomsky, Noam: 1986, *Barriers*, MIT Press, Cambridge.

Chomsky, Noam: 1995, *The Minimalist Program*, MIT Press, Cambridge.

Chomsky, Noam: 2001a, 'Derivation by Phase', in M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, MIT Press, Cambridge, pp. 1–52.

Chomsky, Noam: 2001b, 'Beyond Explanatory Adequacy', ms.

Clark, Cecily: 1958, *The Peterborough Chronicle* (2nd edition 1970), Clarendon Press, Oxford.

Cowling, George H.: 1915, *The Dialect of Hackness*, Cambridge University Press, Cambridge.

Dahl, Östen: 1979. 'Typology of Sentence Negation', *Linguistics* **17**, 79–106.

Davis, Norman: 1971, *Paston Letters and Papers of the Fifteenth Century*, Part I, Clarendon, Oxford.

Dekeyser, Xavier: 1986. 'Relative Markers in the Peterborough Chronicle: 1070–1154', *Folia Linguistica Historica* **7**(1), 93–105.

Demske, Ulrike: 2001, *Merkmale und Relationen*, Mouton de Gruyter, Berlin.

Einenkel, Eugen: 1916, *Geschichte der englischen Sprache*, Strassburg, Karl Truebner Verlag.

Everett, Daniel: 1996, *Why There are No Clitics*, Summer Institute of Linguistics, Arlington.

Fowler, H. W.: 1926 [1950], *A Dictionary of Modern English Usage*, Clarendon Press, Oxford.

Friederici, Angela, Jurgen Weissenborn and Michele Kail: 1991, 'Pronoun Comprehension in Aphasia', *Brain and Language* **41**, 289–310.

Frisch, Stefan: 1997, 'The Change in Negation in ME', *Lingua* **101**, 21–64.

Gelderen, Elly van: 1997, *Verbal Agreement and the Grammar of Its 'Breakdown'*, Niemeyer, Tuebingen.

Gelderen, Elly van: 1998, '*For to* in the History of English', A*merican Journal of Germanic Language and Literature* **10**(1), 45–72.

Gelderen, Elly van: 2002, '*To not* or *not to*' (available at www.public.asu.edu/-gelderen/TO.NOT.doc).

Grossmann, Heinrich: 1906, *Das Angelsächsische Relativ*, Berlin diss.

Haegeman, Liliane: 1995, *The Syntax of Negation*, Cambridge University Press, Cambridge.

Haspelmath, Martin: 1998, 'Does *Grammaticalization need Reanalysis*', *Studies in Language* **22**(2), 315–351.

Heine, Berndt and Mechthild Reh: 1984, *Grammaticalization and Reanalysis in African Languages*, Helmut Buske Verlag, Hamburg.

Henry, Alison: 1995, *Belfast English and Standard English*, Oxford University Press, New York.

Hock, Hans: 1991, *Principles of Historical Linguistics* [second edition], Mouton de Gruyter, Berlin.

Holmberg, Anders: 1991, 'The Distribution of Scandinavian Weak Pronouns', in H. van Riemsdijk and L. Rizzi (eds.), *Clitics and their Hosts*, Europtyp Working Papers II, Tilburg University.

IJbema, Aniek: 2002, *Grammaticalization and Infinitival Complements in Dutch*, Ph.D. Dissertation, University of Leiden.

Jake, Janice: 1994, 'lntrasentential Codeswitching and Pronouns: On the Categorial Status of Functional Elements', *Linguistics* **32**, 271–298.

Jelinek, Eloise: 1984, 'Empty Categories, Case and Configurationality', *Natural Language and Linguistic Theory* **2**, 39–76.

Jelinek, Eloise: 2001, 'Pronouns and Argument Hierarchies', paper presented at the Workshop on Agreement in Argument Structure, Utrecht.

Jespersen, Otto: 1917, *Negation in English and other Languages*, A.F. Høst, Copenhagen.

Jespersen, Otto: 1921, *Language*, Allen and Unwinn, London.

Kean, Marie-Louise: 1979, 'Agrammatism', *Cognition* **7**, 69–84.

Kayne, Richard: 1999, 'Prepositions as Attractors', *Probus* **11**, 39–73.

Kayne, Richard: 2002, 'On Some Prepositions that Look DP-internal', ms.

Kemenade, Ans van: 1999, 'Sentential Negation and Clause Structure in OE', in I. Tieken-Boon et al. (eds.), *Negation in the History of English*, Mouton de Gruyter, Berlin, pp. 147–165.

Kemenade, Ans van: 2002, 'Jespersen's Cycle Revisited', in S. Pintzuk et al., *Diachronic Syntax*, Oxford Univerty Press, Oxford, pp. 51–74.

Kirszner, Laurie and Stephen Mandell: 1992, *The Holt Handbook* [third edition], Harcourt, Brace, Jovanovich, Fort Worth.

Kock, Ernst Albin: 1897, *The English Relative Pronouns, A Critical Essay*, Hjalmar Moeller, Lund.

Kytö, Merja and Matti Rissanen: 1988, 'The Helsinki Corpus of English Texts', in O. lhalainen et al., Rodopi, Amsterdam, pp. 169–179.

Laka, Itziar 1994: *On the Syntax of Negation*, Garland, New York.

Lambrecht, Knut: 1981, *Topic, Antitopic, and Verb Agreement in Non Standard French*, John Benjamins, Amsterdam.

Lightfoot, David: 1979, *Principles of Diachronic Syntax*, Cambridge University Press, Cambridge.

Lyons, Christopher: 1999, *Definiteness*, Cambridge University Press, Cambridge.

MacSwan, Jeff: 1999, *A Minimalist Approach to Intrasentential Code Switching*, Garland, New York.

Mätzner, E.: 1864, *Englische Grammatik*, Weidmann, Berlin.

Mustanoja, Tauno: 1960, *A Middle English Syntax*, Société Néophilologique, Helsinki.

Newmeyer, Frederick: 1998, *Language Form and Language Function*, MIT Press, Cambridge.

Noack, Paul: 1882, *Eine Geschichte der relativen Pronomina in der englischen Sprache*, Calvoer, Goettingen.

Nortier, Jacomine: 1990, *Dutch-Moroccan Arabic Code Switching*, Foris, Dordrecht.

Ordonez, Francisco and Esthela Trevino: 1999, 'Left Dislocated Subjects and the Pro-drop Parameter', *Lingua* **107**, 39–68.

Ouhalla, Jamal: 1990, 'Sentential Negation, Relativised Minimality and the Aspectual Status of Auxiliaries', *The Linguistic Review* **7**, 183–231.

Pierce, Amy: 1994, 'On the Differing Status of Subject Pronouns in French and English Child Language', in B. Lust et al., *Syntactic Theory and First Language Acquisition II*, pp. 319–333.

Quirk, Randolph and Sidney Greenbaum: 1973, *A University Grammar of English*, Longman, London.

Quirk, Randolph and C. L. Wrenn: 1955, *An Old English Grammar*, Methuen, London. [1977 reprint]

Rissanen, Matti: 1999, 'Isn't it? Or is it not?', in I. Tieken-Boon et al. (eds.), *Negation in the History of English*, Mouton de Gruyter, Berlin, pp. 189–205.

Rizzi, Luigi: 1990, *Relativized Minimality*, MIT Press, Cambridge.

Roberts, Ian and Anna Roussou: 1999, 'A Formal Approach to "Grammaticalization"', *Linguistics* **37**, 1011–1041.

Romaine, Suzanne: 1982, *Socio-Historical Linguistics*, Cambridge University Press, Cambridge.

Rosenbach, Anette: 2002, *Genitive Variation in English*, Mouton de Gruyter, Berlin.

Rydén, Mats: 1983, 'The Emergence of *who* as Relativizer', *Studia Linguistica* **37**, 126–134.

Safir, Ken: 1985, *Syntactic Chains*, Cambridge University Press, Cambridge.

Seppänen, Aimo: 2000, 'Preposition-stranding with WH-relatives', *English Language and Linguistics* **4**(2), 295–316.

Simpson, Andrew and Zoe Wu: 2002, 'Agreement Shells and Focus', *Language* **78**(2), 287–313.

Snyder, William: 2000, 'An Experimental Investigation of Syntactic Satiation Effects', *Linguistic Inquiry* **31**(1), 575–582.

Stoelke, H.: 1916, *Die Inkongruenz zwischen Subjekt und Prädikat im Englischen und in den verwandten Sprachen*, Ph.D. dissertation, University of Marburg.

Swan, Michael: 1980, *Practical English Usage*, Oxford University Press, Oxford.

Sweet, Henry: 1871 [1934], *King Alfred's West-Saxon Version of Gregory's Pastoral Care*, Oxford University Press, London.

Taraldsen, Tarald: 1985, 'On Verb-second and the Functional Content', in H. Haider and M. Prinzhorn (eds.), *Verb-Second Phenomena in Germanic Languages*, Foris, Dordrecht, pp. 7–25.

Taraldsen, Tarald: 1992, 'Agreement as Pronoun Incorporation', Paper presented at GLOW conference.

Thorpe, Benjamin: 1861, *Anglo-Saxon Chronicle I and II*, Longman, London.

Traugott, Elizabeth: 1972, *The History of English Syntax*, Holt, Rinehart and Winston, New York.

Traugott, Elizabeth and Berndt Heine: 1991, *Grammaticalization*, John Benjamins, Amsterdam.

Visser, F.: 1963–1973, *An Historical Syntax of the English Language*, 3 Volumes, Brill, Leiden.

Willie, Mary: 1991, *Navajo Pronouns and Obviation*, Ph.D. dissertation, University of Arizona.

Wilson, L.: 1906, *Chaucer's Relative Constructions*, Philological Club of the University of North Carolina.

Wolfram, Walt, Kirk Hazen, and Natalie Schilling-Estes 1999: *Dialect Change and Maintenance on the Outer Banks*, University of Alabama Press, Tuscaloosa.

Wood, Johanna: 1997, *Negation in the Paston Letters*, M.A. thesis, Arizona State University.

Wouden, Ton van der: 1994, *Negative Contexts*, Ph.D. dissertation, University of Groningen.

Zanuttini, Rafaella: 1991, *Syntactic Properties of Sentential Negation*, Ph.D. dissertation, University of Pennsylvania.

Zupitza, Julius: 1959, *Beowulf, Reproduced in Facsimile* [2nd edition, Early English Text Society 245], Oxford University Press, London.