WHERE DID LATE MERGE GO? GRAMMATICALIZATION AS FEATURE ECONOMY

Elly van Gelderen

Abstract. Within early Minimalism, there are Economy Principles such as 'Last Resort', 'Least Effort', and also 'Merge as Late as Possible'. The latter principle makes it possible to explain grammaticalization, an important phenomenon in language change. The general direction of change is one where lexical items are reanalyzed as grammatical (or functional) categories, e.g. verbs as auxiliaries and prepositions as complementizers. One explanation in terms of Late Merge is that if a lexical item is not immediately relevant for theta-structure, it can wait and merge later, rather than merge early with additional movement. In later Minimalism, move is reformulated as internal merge and not considered more uneconomical than regular merge, renamed as external merge. This paper will provide ample examples of the workings of Late Merge and examine how we can capture the grammaticalization effects through an Economy Principle, Feature Economy, that is at work when children acquire a language as well as when language changes.

1. Introduction

This paper shows that the Economy Principle Late Merge (of Chomsky 1995) does not fit in the current version of Minimalism. An alternative Principle is formulated building on the insight that all cross-linguistic variation is dealt with through the positing by the language learner of certain features in the lexicon. Data from language acquisition and language change show that a principle in the internalized grammar of the child (possibly a more general cognitive principle) biases the child towards analyzing certain features as uninterpretable, and also what consequence that has in turn.

The outline is as follows. In section two, I provide some background on the Minimalist framework and on Late Merge as an Economy Principle (see Chomsky 1995: 348). Section three shows, in a very general way, how Late Merge accounts for linguistic changes and for phenomena in language acquisition. Section four goes in more depth and details language change in the CP domain; it turns out that many CP-elements derive from VP-ones through preposing and reanalysis. In section five, I sketch an alternative analysis of Late Merge in terms of Feature Economy, and section six is a conclusion.

2. Minimalism and Economy

In this section, I will provide some background on the Minimalist Program, in particular on Merge. I also give a sample derivation and a discussion of Economy Principles in this model.

2.1. *A Sample Derivation using Merge*

Within a pre-Minimalist system (e.g. Chomsky 1986), a clause consists of an outer layer (the Complementizer Phrase, or CP), an inner layer (the Tense Phrase, or TP), and a thematic layer that contains the verb and its arguments (the traditional Verb Phrase, or VP). The outer layer is responsible for encoding discourse information and for linking one clause to another; the inner layer is involved in the marking of tense and agreement, either through morphology or auxiliaries; and the thematic roles are determined in the lowest layer. Each of these layers can be expanded, and when the sentence is negative a NegP is added. In this framework, syntactic structures are built up using general rules, such as that each phrase consists of a head (X in (1)), and a complement (ZP in (1)) and specifier (YP in (1)):

(1)
$$
\begin{array}{c}
\text{XP} \\
\swarrow \quad \searrow \\
\text{YP} \qquad \text{X}' \\
\qquad \swarrow \quad \searrow \\
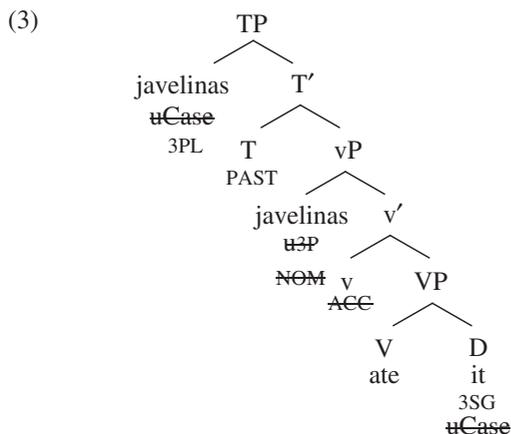\qquad \text{X} \qquad \text{ZP}
\end{array}
$$

This schema holds for both lexical (N, V, A) and grammatical categories (C, D, T).

In the Minimalist Program (Chomsky 1995; 2004; 2007), phrase structure rules are abandoned in favor of a general rule Merge. Merge combines two bundles of features and from Merge, the relations in (1) follow automatically. In this approach, a Modern English derivation proceeds following the following four steps. First, there is a selection of items from the lexicon. Chomsky (2007: 6) suggests the lexicon has "atomic elements, lexical items LI, each a structured array of properties." Abstracting away from features, a lexical array could be {ate, it, T, javelinas}. Secondly, elements are merged, e.g. *ate* and *it* in (2), and one of the two heads projects, in this case V, to a higher VP:

(2)
$$
\begin{array}{c}
\text{VP} \\
\swarrow \quad \searrow \\
\text{V} \qquad \text{D} \\
\text{ate} \qquad \text{it}
\end{array}
$$
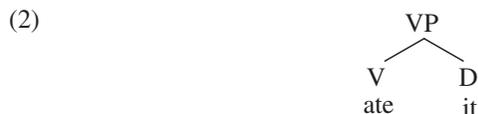
Merge is essential and is the core of the derivational system. Through merge, binary and hierarchical relationships between the merged elements form. We refer to the merged constituents as heads, complements, and specifiers, and to the relations between them via c-command. There is some debate as to whether there is an additional labeling operation and which elements are selected for merge. The labels in (2) are added for convenience only (Chomsky 2004) and it is most likely the theta-requirements that make sure *ate* and *it* merge and not *ate* and PAST. To the VP in (2), a (small) v and subject *javelinas* are merged, as in (3).

Then, after the functional categories such as T (and C) are merged to VP, these heads probe to find a noun with matching features to check agreement and with uninterpretable (u), unvalued Case (indicated as uCase). So, T has interpretable tense features but uninterpretable phi-features (person, number, gender). It probes (i.e. searches) for a nominal it c-commands to agree with. It finds this nominal, or goal, in *javelinas* and each element values its uninterpretable features which then delete. The final structure will look like (3) where the features that are not 'struck through' are interpretable and not subject to elimination. The subject moves to Spec TP, or in other terms is merged from an internal position (Internal Merge) for language-specific reasons:

(3)

```
                        TP
                       /  \
                javelinas   T′
                 u̶C̶a̶s̶e̶    /  \
                  3PL    T    vP
                       PAST  /  \
                        javelinas  v′
                          u̶3̶P̶    /  \
                          NO̶M̶  v    VP
                               A̶C̶C̶  /  \
                                    V    D
                                   ate   it
                                        3SG
                                        u̶C̶a̶s̶e̶
```

In (3), I have also indicated that the light v is a probe providing a value to the case of the object *it*, but the CP layer is not indicated. The derivation in (3) uses early lexical insertion, i.e. a lexicalist approach, as in Chomsky (1995; 2004) but nothing hinges on this.

At some point (in steps/phases), the derivation has to be handed over to the Sensorimotor (SM) and Conceptual-Intensional (CI) systems. The former is responsible for linearization, e.g. what is spelled-out in (3) (see Nunes 2004 for an account of how to decide which copies to spell out). For the CI system, two aspects are relevant, the theta-structure (determined in English through position but in other languages through inherent case) and the discourse information (see Chomsky 2002: 113). There are even two mechanisms responsible for the two, external and internal merge respectively.

2.2. *Economy and Late Merge*

Throughout the different versions of Minimalism, Principles of Economy have played an important role, For instance, 'fewest steps', 'last resort', 'least effort' are all relevant in syntactic derivations (see Chomsky 1995;

Zwart 1996; Collins 1997; Fox 2000). This means that, in building derivations, there are ways to resolve ambiguous structures. "Notice that this approach tends to eliminate the possibility of optionality in derivation. Choice points will be allowable only if the resulting derivations are all minimal in cost" (Chomsky 1995:146) and "economy considerations select among convergent derivations" (1995:348). In the remainder of this section, I outline the Late Merge Principle, as in van Gelderen (2004).

To construct a sentence, we need to select lexical items from the lexicon, put them together, or Merge them, and Move them. In Early Minimalism, Merge "comes 'free' in that it is required in some form for any recursive system" (Chomsky 2004: 108) and is "inescapable" (Chomsky 1995: 316; 378) but Move requires additional assumptions. This means that it is less economical to merge early and then move than to wait as long as possible before merging once. It can be formulated as follows, "all else being equal, wait to merge". This is expressed in (4a). If move is internal merge, we could formulate (4a) as (4b):

(4) **Late Merge Principle (LMP):**
　　a.　Merge as late as possible
　　b.　Avoid Internal Merge.

Late Merge, or Move over Merge, is used in a number of instances. For instance, it is suggested by Chomsky (1995: 348) that Late Merge accounts for the presence of expletive subjects over raising; the principle is used by Fox (2002) to account for Antecedent Contained Deletion and by Bhatt & Pancheva (2004) for the scope of degree clauses. Both Roberts & Roussou (2003) and van Gelderen (2004) use it to account for grammaticalization. The former suggest a change from F^*_{move} to F^*_{merge} (which is parametric) and the latter suggests that, if a lexical item is not relevant to theta-theory, it can merge late.

The question can be asked which lexical items are 'prone' to a reanalysis under the LMP? If non-theta-marked elements can wait to merge outside the VP (Chomsky 1995: 314–5), they will do so. I will therefore argue that if, for instance, a preposition can be analyzed as having fewer semantic features and is less relevant to the argument structure (e.g. *to, for,* and *of* in ModE), it will tend to merge higher (in TP or CP) rather than merge early (in VP) and then move. How does the LMP work in practice? Assuming a lexicalist hypothesis in which a lexical entry "contains three collections of features: phonological ... semantic ..., and formal" (Chomsky 1995: 230), a lexical item such as the light verb *go* might have semantic features of [motion, future, location], as in (5). If *go* occurs in the numeration with another verb, e.g. *get* in (6), one of the semantic features of *go* can be activated, in this case [future] rather than all. In that case, a biclausal structure can be avoided as well:
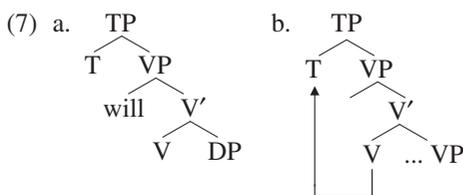
(5) I told Cowslip we were **going** before I left the burrow.

(BNC-EWC 3181)

(6) Anne can HAVE her Mini....Cause I's **gonna** get me a BMW
(http://www.inkycircus.com/jargon/2006/09/anne_can_have_h.html)

Late Merge is thus easy to reformulate in terms of feature Economy, as I will argue in section 4. First, let's look at some more examples.

3. Late Merge at work in change and acquisition

Principle (4) works most clearly in the case of heads. Thus, under Late Merge, the preferred structure would be (7a) with the auxiliary base generated in T, rather than (7b) with the auxiliary in a lower position and moving to T to do double duty, so to speak:

(7) a.
```
        TP
       /  \
      T    VP
    will   / \
          V'
         /  \
        V    DP
```
b.
```
        TP
       /  \
      T    VP
    ↑      / \
          V'
         /  \
        V  ... VP
```

In Old English, the verb *willan* 'want' has semantic features such as [volition, future]. The selection from the lexicon for a sentence as in (7) might include {he, will, think, that}, using Modern English equivalents, and could have (7b) as the result. One of the features of *will* could be reanalyzed as a grammatical tense feature with (7a) as the result. Similar cases of reanalysis are given in (8) and (9), where the *have* and *to* in the (b)-sentences have fewer semantic features:

(8) a. I **have** a garden.
b. I **have** seen the garden.

(9) a. The highway **to** Tucson.
b. I didn't expect **to** think that.

The LMP accounts for the change from lexical to functional head or from functional to higher functional head so frequently described in the grammaticalization literature (e.g. Heine & Kuteva 2002). Other examples are given in Table 1.

Table 1. Examples of the LMP

On, from P to ASP	VP Adverbials > TP/CP Adverbials
Like, from P > C (*like I said*)	Negative objects to negative markers
Modals: v > ASP > T	Negative verbs to auxiliaries
To: P > ASP > M > C	PP > C (*for him to do that ...*)

If economy principles are part of the innate cognitive endowment of the child, it should be possible to see the effects of the LMP in language acquisition. Here I will show that it is. As is well-known, children first acquire lexical and then grammatical/functional categories (e.g. Friederici 1983). Their extension from lexical to functional is a Late Merge phenomenon and I start by providing an illustration of this in Swedish. I then continue with evidence from the CHILDES corpus that English acquisition is similar. I also examine the caregiver speech and show that with the preposition *like* being extended as complementizer, there is no similar pattern to copy in the children's input (at least the one we see), and hence, Late Merge effects must be caused by a principle in the grammar of the child.

Josefsson & Håkansson (2000: 398) show that Swedish "children first acquire the PP and then, directly after that the subordinate clause". They divide the acquisition into three stages, the first one with no prepositions; the second stage with an occasional preposition; the third stage with regular prepositions followed by complementizers. This latter stage is reached for Embla when she is two, with *som* a preposition in (10), introducing a nominal, and a complementizer in (11), introducing a clause:

(10) *precis **som** en ka / **som** en kanin*
    just   like a rab / like a rabbit

(11) *grisen, den **som** heter   Ola*
    the-pig that who is-called Ola
                    (Embla, 27 months, both from Josefsson 2000: 410)

Josefsson & Håkansson say that "most often, the children do not start using complementizers at all until they have reached a 90% use of prepositions" in obligatory contexts.

There is evidence from English, using the Kuczaj corpus (see Kuczaj 1976, and http://childes.psy.cmu.edu/data/Eng-USA), that English *like* and *for* show the same development. The first stages involve the use of *like* as a lexical category, a verb in (12), and a preposition in (13). Abe's first (recorded) example of a *like* complementizer is in (14) and then less than a month later, Abe produces (15):

(12) **like** a cookie   (Abe, 3.7.5)

(13) no the monster crashed the planes down **like** this **like** that
                                    (Abe, 3.7.5)

(14) watch it walks **like** a person walks.   (Abe, 4.9.19)

(15) Daddy # do you teach **like** yaou do [//] **like** how they do in your school?   (Abe, 4.10.1)

Thus, the child 'generalizes' from preposition to complementizer, in a way very similar to what happens in language change. This is all the more interesting since there is no direct evidence for this in the caregivers' speech, who only use *like* after the copulas *look* or *sound*, as in (16), not as a complementizer introducing adverbials, as in (14) and (15):

(16) it looks **like** some birds have eaten some of the bread.

<div align="right">(Kuczaj file 206)</div>

The data with *for* are similar except that here the caregivers do employ *for* as a preposition and a complementizer. The child starts out using *for* as a preposition in (17) to (19), and (20) is the first use as complementizer, followed by many more, e.g. (21) and (22):

(17) Mom # this white one **for** me?   (Abe 2.7.18)

(18) Dad # how come some people have cookies **for** lunch sometimes?

<div align="right">(Abe 3.7, 5)</div>

(19) ok then we could go way # way # way down from the stairs and dig **for** that rock I saw   (Abe 3.7, 5)

(20) yeah and I said I was waiting and waiting **for** you to come and I [/]   (Abe, 3.2.1)

(21) yeah maybe it's time **for** it to rain we'll have a storm. (Abe, 3.6.26)

(22) it's not too high up # but I'm waiting **for** Silver to get ready.

<div align="right">(Abe, 5.0, 11)</div>

In section three, I have given more examples of the Late Merge Principle. I'll now turn to two very specific cases in the history of English and provide evidence for the exact steps with which the change went.

4. Language Change in the CP domain

In this section, I show that both *after* and *for* are typically part of PPs functioning inside the VP but that by Late Old English, they are increasingly preposed and later reanalyzed in that position.

As part of a preposition phrase, *after* functions as a VP-adverbial of time inside the Old English VP, as in (23):[1]

(23) *Fand þa   ðær   inn æþelinga* gedriht *swefan* ***æfter symble***
found then there in   noble   company sleeping after feast
'He found therein a company of nobles sleeping after their feast'

<div align="right">(Beowulf 118–9)</div>

---

[1] The sentences taken from the OED will just be given with their year of appearance. For the other sentences, I have used the usual editions, e.g. Klaeber for *Beowulf*, Thorpe for the *Chronicles*, Skeat for Lindisfarne, and Brook & Leslie for Layamon.

The use of *after* as a complementizer develops through the fronting of the temporal PP headed by *after*, as in (24), and when the object of the preposition becomes a 'bland' demonstrative, as in (25):

(24) a. *Her Leo se æþela papa & se halga forþferde, & **æfter him** Stephanus feng to rice.*
'In this year, Leo the noble and holy pope died and after him, Stephen started to rule'   (Chron A, anno 814 [816])

   b. *& þær wearþ Heahmund biscep ofslægen, & fela godra monna; & **æfter þissum gefeohte** cuom micel sumorlida.*
'And there was Bishop H. killed and many good men, and after this fight came many summer troops'.   (Chron A, anno 871)

(25) a. *Her forðferde Wulfstan diacon on Cilda mæssedæge 7 **æfter þon** forðferde Gyric mæsse preost.*
'In this year died Wulfstan ... and after that died Gyric the priest'.   (Chronicle A, entry for the year 963)

   b. *[Æfter þysan] com Thomas to Cantwarebyri*
'After this, Thomas came to Canterbury.'
(*Chronicle A*, anno 1070)

After this fronting, the temporal PP could be reanalyzed as a clause linker and the clause to which it belonged as an embedded adverbial clause.

The gradual change towards higher base generation of the PP can be shown by comparing the parts of the *Anglo-Saxon Chronicle* done by different scribes in different time periods. The percentages of fronting (which sets the stage for Late Merge) in two different stages, are given in Table 2; the numbers of non-descript prepositional objects is also given.[2]

There is a period where *after that* conjoins clauses, as in (26), but after 1360, *after* is a complementizer on its own, introducing a temporal adverbial clause, as in (27ab). It is reanalyzed as a head (see van Gelderen 2004 for the Head Preference Principle):

(26) **After that** Raleigh had Intelligence that Cobham had accused him, he endeavour'd to have Intelligence from Cobham (HC, EModE2)

Table 2. Percentages of PP fronting and of demonstrative objects (Dem) with *after* in *Chronicle A*.

	Before the year 892	After 893
Fronting	7/26 = 27%	12/22 = 55%
Dem	2/26 = 8%	17/22 = 77%

[2] The *Anglo Saxon Chronicle* describes the years from Julius Caesar's invasion to the years around and after the Norman Invasion of 1066. *Chronicle A* goes to 1070 and is written by one scribe 'Hand I' up to 892. After 892, there is a variety of hands.

Table 3. Grammaticalization of *after*

a.	PP	~~PP~~	900 (Chronicle A) - present
b.	PP (*that*)		950 (Lindisfarne)–1600 (OED 1587)
c.	P *that*		1220 (Lambeth)–1600 (OED 1611)
d.	C		1360 (Wycliff)–present

(27) a. **Aftir** he hadde take þe hooli Goost

(c1360 Wyclif *De Dot. Eccl.* 22).

   b. **After** thei han slayn them   (1366 Mandeville 174).

The stages from VP-adverbial, to topicalized adverbial, to reanalyzed higher adverbial, to complementizer are represented in Table 3.

   I will now turn to *for*, where the same development took place. In Old English, *for* is a preposition indicating location, time, and causation, as in (28) to (30) respectively:

(28) *hlynode*   **for**   **hlawe**
   made-noise before mound
   'It made noise before/around the gravehill'   (Beowulf 1120).

(29) *ða*   *geworden aron in iuh* **for long** *in asca* *... dydon l*
   which become are in you for long in ashes ... did   and
      worhton
      made
   'which were done in you (they) long ago would have repented'

(Lindisfarne Matthew 11.21).

(30) *wen*   *ic þæt ge* **for wlenco** *nalles for wræcsiðum. ac* **for**
   expect I that you for daring not   for misery/exile but for
      higeþrymmum   *Hroðgar sohton.*
      greatness-of-heart Hrothgar sought
   'I expect you were seeking Hrothgar because of your daring and greatness of heart rather than because of being exiled'

(Beowulf 338–9)

The PPs headed by *for* are very often fronted, as in (31) and (32):

(31) **for** *werefyhtum ... ond* **for** *arstafum*   *usic sohtest*
   for fighting   ... and for support (you) us   sought
   'You wanted us to help fight'   (*Beowulf* 457–8).

(32) *ouþer* **for untrumnisse** *ouþer* **for lauerdes neode** *ouþer* **for haueleste**
   ouþer for hwilces cinnes oþer neod *he ne muge þær cumon*
   'either from infirmity or from his lord's need or from lack of means or from need of any other kind he cannot go there'

(*Chronicle E*, anno 675).

As in the case of *after*, they appear with demonstratives, as in (33):

(33) **forþam** *Trumbriht wæs adon of þam biscopdome*
    'because T had been deprived of his biscopric'
                                             (*Chronicle E,* anno 685).

They are reanalyzed as complementizer heads at the end of the Old English period, as in (34) and (35):

(34) **for** *þæt ilc gær warth þe king ded*
    because (in) that same year was the king dead (*Chronicle E*, 1135, 6)

(35) *Locrin 7 Camber to þon scipen comen.* **for** *to habben al þa æhte*
    Locrin and Camber to the ships came for to have all the goods
                                         (Layamon, 1113–4).

In this section, I have illustrated some Late Merge effects when PPs are reanalyzed in higher positions. There are literally hundreds such reanalyses known (see e.g. Heine & Kuteva 2002). The Late Merge Economy Principle would have accounted for this (if a PP moves from VP to CP, reanalyze it in the CP) but since it is no longer available in the late Minimalist Program, I turn to an alternative in the next section.

5. Later Minimalism

Chomsky (2004) uses the terms 'external merge' for the initial merge (the traditional merge) and 'internal merge' when an element is merged for a second time (the traditional move). Since move is seen as a remerge, it is not considered less economical than merge (see Chomsky 2005: 12). One could argue that (4) is still valid, perhaps in the shape of (4b), since internal merge requires steps additional to the ones external merge requires. Traces are no longer allowed, since they would introduce new material into the derivation after the initial selection, and therefore copies are included in the derivation, see (3) above. Internal merge is 'copy, merge, and delete'. Since the derivation contains more copies of the lexical item to be internally merged, and since those copies have to be deleted, (4) could still hold as an Economy Principle. In addition, Chomsky (2005: 14) suggests that a real difference between the two kinds: external merge is relevant to the argument structure, whereas internal merge is relevant for scope and discourse phenomena. This indicates a crucial difference between the two kinds of operations and that difference is expressed in the LMP.

   Above, I mentioned the crucial role of features and of probes. In this model, syntax is inert and the same cross-linguistically. It is the lexical features that differ and change. I therefore reformulate Late Merge in terms of feature change and loss. From Chomsky (1995) on, features are divided in interpretable (relevant at LF) and uninterpretable (not relevant

to the interpretation). Interpretable features are acquired before unin-terpretable ones, as argued in Radford (2000), but are later reinterpreted as uninterpretable ones, triggering the functional/grammatical system. The same happens in language change. For instance, changes from verbs to auxiliaries and from prepositions to complementizers can be accounted for by arguing that their (initially) semantic features are reanalyzed as interpretable ones (iF) and then as uninterpretable ones (uF). For instance, the PP inside a VP would have semantic features of time and these are reanalyzed as grammatical time when the PP is used to link an adverbial clause, as in (24) and (25) above. This can be formalized as in (36):

(36) **Feature Economy**
Minimize the semantic and interpretable features in the derivation, e.g:
VP-Adverbial CP-Adverbial C-Head
semantic >   [iF] >   [uF]

Chomsky (2004; 2007) argues that we need to attribute as little as possible to UG and instead rely as much as possible on principles not specific to the faculty of language. Many Economy Principles, (4) and (36) included, fall into this latter category in that they reduce the computational burden.

A preposition such as *after* has semantic features (e.g. [time, order, past]) and phonological ones (two syllables, etc). The latter are not accessible during the derivation. In addition, there are formal features, which are accessible during the computation and include categorial, Case, and phi-features, at least in Chomsky (1995: 230–2). Assuming that prepositional probes have unvalued phi-features, they value the Case of the DP in their domain. So, the Case of the DP is valued after agreement with an appropriate probe (I use ACC to show this but nothing hinges on this):

(37)
```
            PP
          /    \
        P       DP
      after     him

      [u-phi]   [3S]

      [ACC]     [uCase]
```
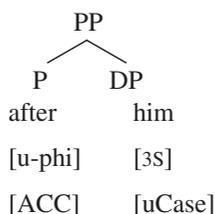
Thus, there is a formal uninterpretable and unvalued feature [u-phi] that makes prepositions into probes. This is the feature that is relevant for the derivation; other features are in fact a burden on the computational system. Language learners and users thus use (36) to eliminate [ACC] from the lexical item. With the interpretable [ACC] feature reanalyzed, the structure will be as in (38), and the same for *like* and *for* (and a number of others). The uninterpretable, unvalued features of C will

probe into the clause it c-commands, and find a goal in the lower TP to value its phi-features. It is well-known that CPs (as subjects) trigger third person singular agreement on the verb. This is expected if the complementizer has phi-features (that are overt in many languages):

(38)

```
            CP
          /    \
        C       TP
      after     [3S]
     [u-phi]
```

The change from (37) to (38) is from having the full topicalized PP function as complementizer to having the P head function as complementizer. What happens to the C in the case of (38)? I assume this C is not specified for grammatical temporality (yet) but that the semantic features of [time] connected to *after* express that, as in (39a). Once the PP is base generated in the CP, [time] is analyzed as an [i-time] on *after*, as in (39b). After the reanalysis of *after* as C in (38), *after* keeps those [i-time] features, or they may be reanalyzed to [u-time] and need another adverb of time with semantic features. A complete picture of the changes is given structurally in (39) and, for both *after* and *for*, in terms of Feature Economy in (40):

(39)   a.

```
              CP                    >        b.              CP                      >
           /     \                                        /      \
        PP        TP                                   PP          C'
       /   \                                          /   \       /   \
      P     DP                                       P     DP    C     TP
    after   him                                    after  that
   [u-phi]  [3S]                                  [u-phi] [3S]
   [ACC]    [uCase]                               [ACC]   [uCase]
   [time]                                         [i-time]
   (=(24))                                        (=(25) and (26))
```

c.

```
              CP
           /     \
          C       TP
        after     [3S]
       [u-phi]
       [i-time]
       (=(27))
```

(40)   P   >   P   >   C

    [u-phi]   [u-phi]   [u-phi]

    [ACC]   [ACC]   [i-time]/[u-time]

    [time]   [i-time]

    (in (24), (31), (32)) (in (25), (26), (33)) (in (27), (34), (35))

Concluding, Late Merge is argued to be a motivating force of linguistic change, accounting for the change from specifier to higher specifier (in the case of the PPs) and head to higher head (in the case of modals). The reason is that these Principles help a child reanalyze their linguistic input. I have reformulated the LMP as a Feature Economy Principle. Feature loss, I argue, can then be responsible for certain grammaticalizations. One could think of feature loss as happening through the selection into the Numeration.

In this paper, I do not touch on other changes that are easily accounted for by (36), namely changes that are known as Jespersen's Cycle where a semantically negative element is reanalyzed as having interpretable features and subsequently uninterpretable features as well as subject pronouns (with interpretable features) being reanalyzed as agreement markers (i.e. as uninterpretable features).

6. Conclusion

I have looked at the grammaticalization that used to be accounted for by the Late Merge Principle, e.g. in Roberts & Roussou (2003) and van Gelderen (2004). I have argued for a reformulation focusing on lexical rather than derivational characteristics. This results in a Feature Economy Principle (from semantic to interpretable to uninterpretable) that accounts readily for the grammaticalization from preposition to complementizer, and other grammaticalizations 'up the tree'. Chomsky (2004; 2007) argues that we need to attribute as little as possible to UG and instead rely as much as possible on principles not specific to the faculty of language, i.e. 'third factor principles'. Many Economy Principles, (36) included, fall into this latter category in that they reduce the computational burden. Each language learner decides on the basis of the language s/he hears which features to include.

References

BHATT, R. & PANCHEVA, R. 2004. Late Merger of Degree Clauses. *Linguistic Inquiry* 35.1, 1–45.
BRITISH NATIONAL CORPUS (BNC). Available online.
BROOK, G. & LESLIE, R. 1963. *Layamon: Brut*. Oxford: Oxford University Press.

CHILD LANGUAGE DATA EXCHANGE SYSTEM (CHILDES)

CHOMSKY, N. 1986. *Barriers*. Cambridge: MIT Press.

CHOMSKY, N. 1995. *The Minimalist Program*. Cambridge: MIT Press.

CHOMSKY, N. 2002. *On Nature and Language*. Cambridge: Cambridge University Press.

CHOMSKY, N. 2004. Beyond Explanatory Adequacy. *Structures and Beyond*, eds. A. Belletti, 104–131. Oxford: Oxford University Press.

CHOMSKY, N. 2005. Three factors in language design. *Linguistic Inquiry* 36.1, 1–22.

CHOMSKY, N. 2007. Approaching UG from below. In Uli Sauerland et al (eds), *Interfaces + Recursion = Language*, 1–29. Berlin: Mouton de Gruyter.

COLLINS, C. 1997. *Local Economy*. Cambridge: MIT Press.

FOX, D. 2002. Antecedent-Contained Deletion and the Copy Theory of Movement. *Linguistic Inquiry* 33.1, 63–96.

FRIEDERICI, A. 1983. Children's sensitivity to function words during sentence comprehension. *Linguistics*, 21, 717–739.

GELDEREN, E. VAN 2004. *Grammaticalization as Economy*. Amsterdam: John Benjamins.

HEINE, B. & KUTEVA, T. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

JOSEFSSON, G. & HÅKANSSON, G. 2000. The PP-CP Parallelism Hypothesis and Language Acquisition. *The Acquisition of Scrambling and Cliticization*, eds. Powers S. & Hamann C., 397–422. Dordrecht: Kluwer.

KLAEBER, F. 1922 [1941]. *Beowulf*. Boston: Heath & Co.

KUCZAJ, S. 1976. *-Ing, -s, -ed: A study of the acquisition of certain verb inflections*. University of Minnesota PhD.

NUNES, J. 2004. *Linearization of chains and sideward movement*. Cambridge: MIT Press.

OXFORD ENGLISH DICTIONARY (OED)

RADFORD, A. 2000. Children in Search of Perfection: Towards a Minimalist Model of Acquisition. *Essex Research Reports in Linguistics* 34. (http://privatewww.essex.ac.uk/~radford/PapersPublications/perfection.htm).

ROBERTS, I. & ROUSSOU A. 2003. *Syntactic Change*. Cambridge: Cambridge University Press.

SKEAT, W. 1881–7. (ed) *The Gospel according to St. Matthew, St. Mark, St. Luke and St. John*. Cambridge.

THORPE, B. 1861. *Anglo-Saxon Chronicle I and II*. London: Longman.

ZWART, J.-W. 1996. 'Shortest Move' versus 'Fewest Steps'. *Minimalist Ideas*, eds. W. Abraham & S.D. Epstein, H. Thráinsson & C.J.-W. Zwart, 305–327. Amsterdam: John Benjamins.

Elly van Gelderen
Department of English
Arizona State University
Box 870302
Tempe, AZ 85287–0302
USA
ellyvangelderen@asu.edu