Where did Late Merge go? Grammaticalization as Feature Economy

Elly van Gelderen

June 2006

Within early Minimalism, there are Economy Principles such as `Last Resort', `Least Effort', and also `Merge as Late as Possible'. The latter principle made 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 was that if a lexical item was not immediately relevant for theta-structure, it could wait and merge later, rather than early merge and additional movement. In later Minimalism, move is reformulated as internal merge and not considered any more uneconomical than regular merge, renamed as external merge. This squib will 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.

Keywords: Economy, Minimalism, Grammaticalization, Complementizers, Prepositions, Features.

1 Introduction

In this section, I present some background on Economy Principles, Late Merge, and Universal Grammar (UG). In section 4, I provide more details on Minimalist features.

Within early Minimalism, there are Economy Principles such as `Last Resort', `Least Effort', and also `Merge as Late as Possible' (e.g. Chomsky 1995; Zwart 1996; and Collins 1997). These Economy Principles give choices, unlike regular UG principles such as c-command which are never violable. Preposition Stranding, for instance, can be seen as due to a `Move as little as possible', but it isn't absolute. Even in spoken English, there is an occasional preposition that has moved along. Some estimates for formal spoken English is that 14% are not stranded in relative clauses.

Late Merge, or Move over Merge, is used in a number of instances. It can be formulated as follows, "all else being equal, wait to merge". 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. Below, in sections two and three, I provide several examples of lexical elements that are initially used quite early in the derivation but then are reanalyzed as also being mergeable later in the derivation.

Later Minimalism (Chomsky e.g. 2004) argues that, due to the Inclusiveness Condition, movement cannot introduce new elements. Traces are therefore abandoned in favor of a copy and delete system and Move is replaced by Internal Merge (or remerge) and not seen as uneconomical. In this squib, I show that there are real Late Merge effects in language acquisition and change and I will argue that the effects of the Late Merge Principle can be realized using the idea (present since Borer 1984) that cross-linguistic variation is in the lexicon and that syntax is inert. If so, all variation is in the lexicon and the difference between a preposition and a complementizer and between a verb and an auxiliary can be seen in terms of features. Feature loss, I argue, can then be responsible for certain grammaticalizations. One could think of feature loss as happening in the numeration, as a `Numeration Sloppiness', or in the lexicon. I will suggest the latter.

The outline is as follows. In section two, I provide examples of what I will continue to call Late Merge in language acquisition, and in section 3 in language change. Section 4 will sketch an alternative analysis of Late Merge in terms of Feature Economy, and section 5 is a conclusion.

2 Acquisition

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, and hence, Late Merge effects must be caused by a principle in the grammar of the child.

Josefsson (2000: 398) shows that Swedish "children first acquire the PP and then, directly after that the subordinate clause". She divides 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, where *som* is a preposition in (1) and a complementizer in (2):

(1) *precis **som** en kan/ **som** en kanin*
just like a rab/ like a rabbit

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

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

In the next part of this section, I will provide evidence 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 (3), and a preposition in (4). Abe's first (recorded) example of a *like* complementizer is in (5) and then less than a month later, Abe produces (6):

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

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

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

(6) Daddy # do you teach **like** you 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 as we'll see in the next section. 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 (7), not as a complementizer introducing adverbials, as in (5) and (6):

(7) it looks **like** some birds have eaten some of the bread. (Kuczaj file 206)

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 (8) to (10), and (11) is the first use as complementizer, followed by many more, e.g. (12) and (13):

(8) Mom # this white one **for** me? (Abe 2.7.18)
(9) Dad # how come some people have cookies **for** lunch sometimes? (Abe 3.7, 5)
(10) ok then we could go way # way # way down from the stairs and dig **for** that rock I saw (Abe 3.7, 5)
(11) yeah and I said I was waiting and waiting **for** you to come and I [/] (Abe, 3.2.1)
(12) yeah maybe it's time **for** it to rain we'll have a storm. (Abe, 3.6.26)
(13) it's not too high up # but I'm waiting **for** Silver to get ready. (Abe, 5.0, 11)

An interesting difference between *like* and *for* is that (5) and (6) are the only two instances of *like* as a complementizer that occur in the data, whereas *for* occurs very frequently (21 times for Abe; and 35 for the adults he interacts with).

3 Language Change in the CP domain

In this section, I first examine the origin of the `high' adverbs (as in Cinque 1999) and then examine the complementizers *after* and *for*. Their historical development provides evidence for Late Merge.

3.1 High Adverbs

Cinque (1999) formulates his (by now well-known) functional hierarchy and the highest adverbs are *frankly*, *fortunately*, *allegedly*, and *probably*. He terms them speech act,

evaluative, evidential, and epistemic respectively. In many languages, the higher adverbs are not base generated in high positions but moved from lower positions, and are sometimes separate clauses in which the adverb modifies the verb of the subordinate clause, e.g. in the Dutch (14) and Bulgarian (15):

(14) [*Om het eerlijk te zeggen] begrijp-ik er niets van*
for it honestly to say understand-I there nothing of
`Honestly, I don't get it'.

(15) [*Chestno kazano] nishto ne razbiram*
Frankly spoken nothing not understand-1S.PRES
'Frankly, I don't understand anything.' (Mariana Bahtchevanova p.c.)

The next step will be for the adverbs *eerlijk* and *chestno* to be reanalyzed as CP adverbs.

In English, the high adverbs are all loans and fairly recent CP adverbs. Thus, the data on English *(un)fortunately* shows Late Merge effects since these adverbs are reanalyzed from VP/vP to CP-adverb. The actual word is borrowed from French, and first appears as an adjective, noun, and verb, and as a VP/vP-adverb, as in (16) and (17), meaning `in an (un)happy, (un)lucky, (un)successful manner, done by an agent' <1>. (Note that the *-ly* ending does not appear till later). In the 17th century, the adverb is part of a non-finite clause, as in (18), and becomes a CP-adverb in the 18th century, as in (19):

(16) Whan a man..clymbeth vp and wexeth **fortunat**. (OED, 1386)

(17) After this victorye **fortunately** obteined (OED, 1548)

(18) but in the fight loosing most of his men, himself with a few escap'd: only Theobald the Kings brother, and the whole wing which he commanded, [**unfortunately** cut off], made the Victory to Ethelfrid less intire. (1670, Milton, History X, 146)

(19) **Fortunately**, Lord De la War..met them the day after they had sailed (OED, 1796)

3.2 After and for: from P to C

The preposition (and adverb) *after* has always been in the English language. Its use as a complementizer is more recent. The way it developed is that first the PP headed by *after* was fronted, as in (20), and the object of the preposition became a `bland' demonstrative, as in (21):

(20) a. [**æfter him**] Stephanus feng to rice.
`after him (i.e. Pope Leo), Stephanus became pope'.
(*Chronicle A*, anno 814 [816])

b. [**æfter þissum gefeohte**] cuom micel sumorlida.
`after this fight, there came a large summer-force'.
(*Chronicle A*, anno 871)

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

b. [**æfter ðon**] uutedlice ic eftariso ic forlioro vel iowih in galileam
`after that, surely I arise-again I come before you in Galilee'
(*Lindisfarne Gospel, Matthew* 26. 32).

The change towards gradual higher base generation can be shown by comparing the parts of the *Anglo-Saxon Chronicle* done by different scribes. The percentages of fronting (which sets the stage for Late Merge) in two different stages, are given in Table 1; the numbers of non-descript prepositional objects is also given <2>.

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

Table 1: Percentages of PP fronting and of demonstrative objects (Dem) with *after* in *Chronicle A*.

There is a period where *after that* is the complementizer, as in (22) and (23), but after 1360, *after* is a complementizer on its own, as in (24) and (25), reanalyzed as a head. The stages are represented in Table 2:

(22) **After that** the king hadde brent the volum
(Wyclyf 1382, taken over in Coverdale 1535 and KJV 1611, from the OED).

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

(24) **Aftir** he hadde take þe hooli Goost (c1360 Wyclif *De Dot. Eccl.* 22).

(25) **After** thei han slayn them (1366 Mandeville174).

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

Table 2: Grammaticalization of *after*

I will now turn to *for*, where the same development took place. In Old English, *for* is a preposition indicating causation, as in (26) and (27). The PPs headed by *for* are very often fronted and are complementizers in (28) and reanalyzed as complementizer heads at the end of the Old English period, as in (29) and (30):

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

(27) *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).

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

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

(30) *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 VP adverbs and 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 but since it is no longer available, I turn to an alternative in the next section.

4 Feature Grammaticalization

Three kinds of features are seen as relevant, namely phi-features on the Probe, structural Case on the Goal, and EPP/OCC on the Probe (see Chomsky 2004: 116). Each language learner decides on the basis of the language s/he hears which features to include. Using these features, a derivation proceeds as follows. Lexical Items are selected (as a lexical array) from the lexicon to be accessed in the Derivation. Merge then takes two items and puts them together, initially through External Merge (the vP shell).

After probes such as T and v are merged, these probes examine their c-command domains, and Agree with the closest DP. This operation values their unvalued phi-features, and in turn values the Case on the DP, as in the simplified (31). This valuation is indicated by `strike through':

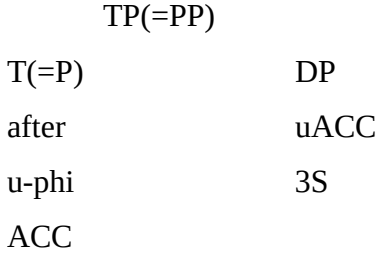
(31)

CP							
C		TP					
.		T'					
	T		vP				
	Pres	they		v'			
	u~~3P~~	~~uCase~~	v		VP		
	~~Nom~~	3P	~~Acc~~	V		D	
				see		it	
						3S	
						~~Acc~~	

The EPP/OCC feature ensures internal merge to certain positions, but is not relevant to this paper and hence, it is not indicated in (31).

Having given some background on features, we can now proceed to reformulate Late Merge in terms of feature loss. A preposition such as *after* has semantic features (e.g. [time, order, past]) and phonological ones (two syllables, etc), not accessible during the Derivation, as well as formal features, accessible during the computation. In Chomsky (1995: 230-2), the formal features include categorial, Case, and phi-features. In later work, following Marantz (1997), lexical items are seen as not specified for category but as roots that are nominalized or verbalized through Merge. I assume, adapting Pesetsky & Torrego (2004), that prepositions are in fact special Ts. This means they have unvalued phi-features and value the Case of the DP in their domain:

(32)



Thus, there is a formal uninterpretable and unvalued feature 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 (33) to eliminate [ACC] from the lexical item:

(33) **Economy of Features**
Minimize the interpretable features in the derivation

With the interpretable feature removed, the structure will be as in (34), and the same for *like* and *for* (and a number of others). The uninterpretable, unvalued features of T 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):

(34) TP(=CP)

T(=C) TP

after 3S

u-phi

So far, (33) accounts for grammaticalizations of prepositions to complementizers. How about the change of VP/vP-adverbials to CP-ones? If the light v is connected to agentivity, the manner adverb would be connected to this layer as well since it is impossible to do something in a fortunate or frank manner if there is no agent. The CP layer marks discourse functions, e.g. the intent of the speaker, rather than the agent, and the grammaticalization from VP/vP to CP-adverb involves a reanalysis from agent-oriented (theta-related) to speaker-oriented (discourse-related). The mechanism of change is clear, namely the VP/vP-adverb is often preposed and then reanalyzable as CP-adverb. The reasons for this preposing are not so clear (possibly main clauses need to indicate discourse functions) and a precise account for CP-adverbs therefore remains for future research.

5 Conclusion

In this paper, I have looked at two kinds of grammaticalizations that used to be discussed as cases of Late Merge, 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 that accounts readily for the grammaticalization from preposition to complementizer. A fairly similar reanalysis from VP/vP to CP-adverb is harder to account for since it is not quite clear what features adverbs check, if any.

Chomsky (2004; 2006: 2-3) 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, (33) included, fall into this latter category in that they reduce the computational burden.

Abbreviations and notes

CHILDES	Child Language Data Exchange System
OED	Oxford English Dictionary
S	singular
u	uninterpretable
UG	Universal Grammar
1P	first plural etc.

<1> The sentences that I have found using 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.

<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.

References

Bhatt, Rajesh & Roumyana Pancheva 2004. "Late Merger of Degree Clauses". *Linguistic Inquiry* 35.1: 1-45.

Borer, Hagit 1984. *Parametric Syntax*. Dordrecht: Foris.

Brook, G. & R. Leslie 1963. *Layamon: Brut*. Oxford: Oxford University Press.

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

Chomsky, Noam 2004. "Beyond Explanatory Adequacy". In Adriana Belletti (ed.). *Structures and Beyond*, 104-131. Oxford: Oxford University Press.

Chomsky, Noam 2006. "Approaching UG from below". Ms.

Cinque, Guglielmo 1999. *Adverbs and Functional Heads*. Oxford: OUP.

Collins, Chris 1997. *Local Economy*. Cambridge: MIT Press.

Fox, Danny 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, Elly van 2004. *Grammaticalization as Economy*. Amsterdam: John Benjamins.

Heine, Bernd & Tania Kuteva 2002. *World Lexicon of Grammaticalization.* Cambridge: Cambridge University Press.

Josefsson, Gunløg 2000. "The PP-CP Parallelism Hypothesis and Language Acquisition". In S. Powers et al. (eds) *The Acquisition of Scrambling and Cliticization*, 397-422. Dordrecht: Kluwer.

Klaeber, Fr. 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.

Marantz, Alex "No Escape from Syntax". ms.

Pesetsky, David & Esther Torrego 2004. "Tense, Case, and the nature of Syntactic Categories". ms.

Roberts, Ian & Anna Roussou 2003. *Syntactic Change*. Cambridge: Cambridge University Press.

Skeat, Walter 1881-7. (ed) *The Gospel according to St. Matthew, St. Mark, St. Luke and St. John*. Cambridge.

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

Zwart, Jan-Wouter 1996 "`Shortest Move' versus `Fewest Steps'". In Werner Abraham, et al (eds). *Minimal Ideas*, 305-327. Amsterdam: John Benjamins.