

**Economy of Merge and Grammaticalization:
Two steps in the Evolution of Language**

Elly van Gelderen

6 September 2006

Grammaticalization is an easily observable process in the history of languages and has therefore frequently been seen as involved in the evolution of language from its earliest stage to the present. If, as has been argued by Givón (1979) and Bickerton (1990), the proto-language had an emphasis on the pragmatic rather than the syntactic function, we expect grammaticalization, which takes linguistic expressions that are pragmatically relevant (e.g. as Topics) and incorporates them into the syntax (e.g. as Subjects), to be crucial in language evolution. This syntactic incorporation cannot be done without a process linking two elements. The latter process is therefore extremely important and was, according to Chomsky, the "Great Leap Forward" in the evolution of humans" (Chomsky 2005: 11). A slight rewiring of the brain might have made the operation Merge possible and, in its turn, Merge made syntax possible by combining concepts into multiple unit expressions, with in principle infinite recursion.

This paper has a relatively narrow focus in that it will examine the consequences of Merge for language evolution. I argue that there are two crucial steps to the evolution of syntax, namely Merge and, following from principles connected with it, Grammaticalization. The emergence of Merge brings with it certain relations such as Specifiers, Heads, Complements, and c-command. Heads, complements, and specifiers in turn define argument structure (e.g. as in Hale & Keyser 2002). In addition, I argue that there was a second development due to Principles of Merge Economy. These were responsible for processes known as grammaticalization. For instance, even though subordination was in principle possible, it probably didn't arise till the grammaticalization of complementizers. Merge brought about the first step of linguistic evolution but Principles connected with it were responsible for further language evolution. These processes continue up to the present.

The outline is as follows. In section 1, I discuss Merge and Phrase Structure rules and their relevance to linguistic evolution. In section 2, I discuss Grammaticalization and

the Economy Principles that account for it, and provide examples of the linguistic cycle. In section 3, I propose a link between Evolution and Grammaticalization, and in section 4, a Principle is suggested that incorporates new lexical items. In section 5, I provide a conclusion and look into some criticisms of the view that grammaticalization provides us insight into language evolution.

1 Universal Grammar, Merge, and its implications

In this section, I discuss the generative model and in particular Universal Grammar, how Merge works in a derivation, and also what further structural relationships it is responsible for. I then comment on language evolution.

Starting in the 1950s, Chomsky and the generative model he develops present an alternative to then current behaviorist and structuralist frameworks. Chomsky focuses not on the structures present in the language/outside world but on the mind of a language learner/user. The input to language learning is seen as poor (the 'poverty of the stimulus' argument) since speakers know so much more than what they have evidence for in their input. How do we know so much on the basis of such impoverished data? The answer to this problem, Plato's problem in Chomsky (1986), is Universal Grammar (hence UG), the initial state of the faculty of language, a biologically innate organ. UG helps the learner make sense of the data and build up an internal grammar. Initially, many principles were attributed to UG but currently (e.g. Chomsky 2004; 2006), there is an emphasis on principles not specific to the faculty of language, i.e. UG, but to "general properties of organic systems" (Chomsky 2004: 105). Merge is one such operation that can be seen as a UG principle (Chomsky 2006: 4) but also as one possibly "appropriated from other systems" (Chomsky 2006: 5). I'll now turn to how a sentence is actually produced.

In the Minimalist Program, the most recent generative framework (Chomsky 1995; 2004), a Modern English derivation proceeds as follows. Merge combines two items, e.g. *see* and *it* in (1), and one of the two heads projects, in this case V, to a higher VP:

(1) VP
 V D
 see it

There is some debate as to the labels, which for most people are added for convenience only. I will also add them. Phrase structures are built using merge and move (also called external and internal merge respectively). The VP domain is the thematic-layer, i.e. where the argument structure is determined.

Apart from merge, there are "atomic elements, lexical items LI, each a structured array of properties (*features*)" (Chomsky 2006: 4). Each language learner selects the features compatible with the input. The main kinds involve Case, agreement (also known as phi-features), and displacement to subject position. Features come in two kinds. Interpretable ones include number on nouns and are relevant at the Conceptual-Intentional interface. Uninterpretable features include agreement features on verbs and Case features on nouns. These features need to be valued and deleted. Continuing the derivation in (1) will make the function of features clearer.

After adding a (small) *v* and subject *they* to (1), functional categories such as T (and C) are merged to VP. Agree ensures that features in TP (and CP) find a noun or verb with matching (active) features to check agreement and Case. So, T has interpretable tense features but uninterpretable phi-features. It probes ('looks down the tree') for a nominal it c-commands to agree with. It finds this goal in *they* and each element values its uninterpretable features: the noun's Case as nominative and the verb's phi-features as third person plural. The final structure will look like (2) where the features that are not 'struck through' are interpretable. The subject moves to Spec TP for language-specific reasons:

(2) TP
 They T'
 ~~uCase~~ T vP
 3P Pres ~~they~~ v'
 ~~u3P~~ v VP
 Nom Acc V D
 see it

and a Specifier, resulting in a locative structure; in (4c), there is a Head and a Specifier, the traditional adjectival predicate, and in (4d), there is just a Head (examples are from Hale & Keyser 2002, with XP, YP, and ZP arbitrary phrases):

- | | | | | | | |
|-----|----|-----|-------|----|---------|-------|
| (4) | a. | X | | b. | X | |
| | | X | Y | | ZP | X |
| | | | laugh | | book | X |
| | | | | | with/on | Y |
| | | | | | | shelf |
| | c. | Y | | d. | X | |
| | | ZP | Y | | sky | |
| | | sky | Y | | X | |
| | | | | | clear | |

So, the first step in the evolution of syntax is Merge. It brings with it notions of headedness (once you merge two elements, one determines the resulting label) and binarity. These notions also determine possible argument structures. The next step is for grammatical heads, such as auxiliaries and prepositions, to appear, as we will discuss in section 3. First, however, we need to look at the mechanisms of adding grammatical elements. I will argue that Economy of Merge is responsible for that.

2 Grammaticalization as Economy

As is well-known, grammaticalization is a process whereby lexical items lose phonological weight and semantic specificity and gain grammatical functions. Grammaticalization has frequently been investigated in a functionalist framework. Recently, however, structural accounts have started to appear (e.g. Abraham 1993; Roberts & Roussou 2003; van Gelderen 2004) accounting for the cyclicity of the changes involved. Van Gelderen, for instance, uses Economy Principles that help the learner acquire a grammar that is more economical, and as a side-effect more grammaticalized.

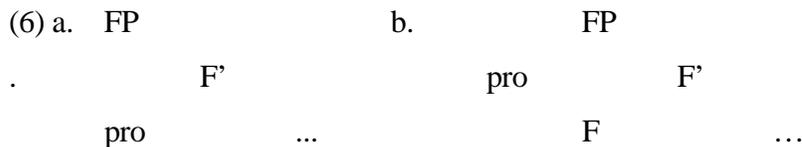
Two Economy Principles, provided as (5) and (13) below, are formulated in van Gelderen (2004). They are part of UG and help learners construct a grammar. They are similar to principles such as c-command, in that they remain active in the internalized grammar and therefore also aid speakers in constructing sentences. They are different from absolute principles such as c-command because prescriptive and innovative tendencies can counteract them.

Principle (5) is a principle at work in the internalized grammar and holds for merge (projection) as well as move (checking). It is most likely not a principle specific to language but a property of organic systems:

(5) **Head Preference Principle (HPP):**

Be a head, rather than a phrase.

This means that a speaker will prefer to build structures such as (6a) rather than (6b). The FP stands for any functional category and the pro(noun) is merged in the head position in (6a), and in the specifier position in (6b). Other categories, such as adverb or preposition, work the same way:



The speaker will only use (b) for structures where a phrase is necessary, e.g. coordinates. There may also be prescriptive rules stopping this change (as there are in French, see Lambrecht 1981).

As is well-known, native speakers of English (and other languages) producing relative clauses, prefer to use the head of the CP (the complementizer *that*) rather than the specifier (the relative pronoun *who*) by a ratio of 9:1 in speech. As expected, children acquiring their language obey this same economy principle. Thus, according to Diessel (2004), young children produce only stranded constructions in English, as in (7):

(7) those little things that you play with (Adam 4:10, from Diessel 2004: 137).

Once they become (young) adults, they are taught to take the preposition along.

The Head Preference Principle is relevant to a number of historical changes: whenever possible, a word is seen as a head rather than a phrase. In early Germanic, the negative element *ne* precedes the verb (as in other Indo-European languages). In the North Germanic languages, this *ne* is phonologically very weak. As Wessén (1970: 100) puts it "[d]a die Negation schwachtonig war, machte sich das Bedürfnis nach Verstärkung stark geltend". This strengthening comes in the form of an enclitic *-gi* that attaches to regular words. This results in *eigi* 'not', as in (8), *aldrigi* 'never', *eitgi/ekki* 'nothing', and numerous other forms:

(8) *Pat mæli ek eigi* Old Norse
that say-1S I not
'I am not saying that' (from Faarlund 2004: 225).

Faarlund (2004: 225) states that the *-gi* suffix is no longer productive in Old Norse but rather that it is part of negative words. That means that *eigi* and other negatives in Old Norse are phrasal adverbs, in specifier positions, as is obvious because they trigger V-second, as in (9):

(9) *eigi vil ek Pat* Old Norse
not want I that
'I don't want that' (Faarlund 2004: 225).

For Modern Norwegian, Bondi Johannessen (1997; 2000) argues that *ikke* 'not' (derived from *eigi*) is a head. That means that between Old Norse and Modern Norwegian, the negative is reanalyzed from specifier to head. That it is a head is also clear from (10) since it adjoins to the verb. An expected further change is that the head weakens phonologically and this is indeed the case as is fairly obvious from sentences such as (10), pretty common according to native speakers:

- (10) *Men detta æ'kke et forslag som vi har interesse av* Norwegian
but that is-not a proposal that we have interest in
'But that's not a proposal we are interested in' (from Solstad 1977: 70).

This development is similar to that in English with negative auxiliaries such as *don't*.

The next stage after (10) is when the weakened negative is reinforced. This may be occurring in certain varieties of Norwegian. Thus, Sollid (2002) argues that in the Northern Norwegian dialect of Sappen a double negative is starting to occur, as in (11):

- (11) *Eg har ikke aldri smakt sånne brød* Sappen Norwegian
I have not never tasted such bread
'I haven't ever tasted that kind of bread' (Sollid 2002).

She argues this is under the influence of Finnish, which may well be the case. This would, however, not be possible if the grammar wasn't ready for this, i.e. if *ikke* weren't already a head.

The changes can be summarized in Figure 1, where (a) and (b) represent Old Norse, (c) is Norwegian, and (d) represents a variety such as Sappen with the verb moving through the Neg head and the negative being reinforced by a new specifier. Traditionally, these changes are known as Jespersen's Cycle.

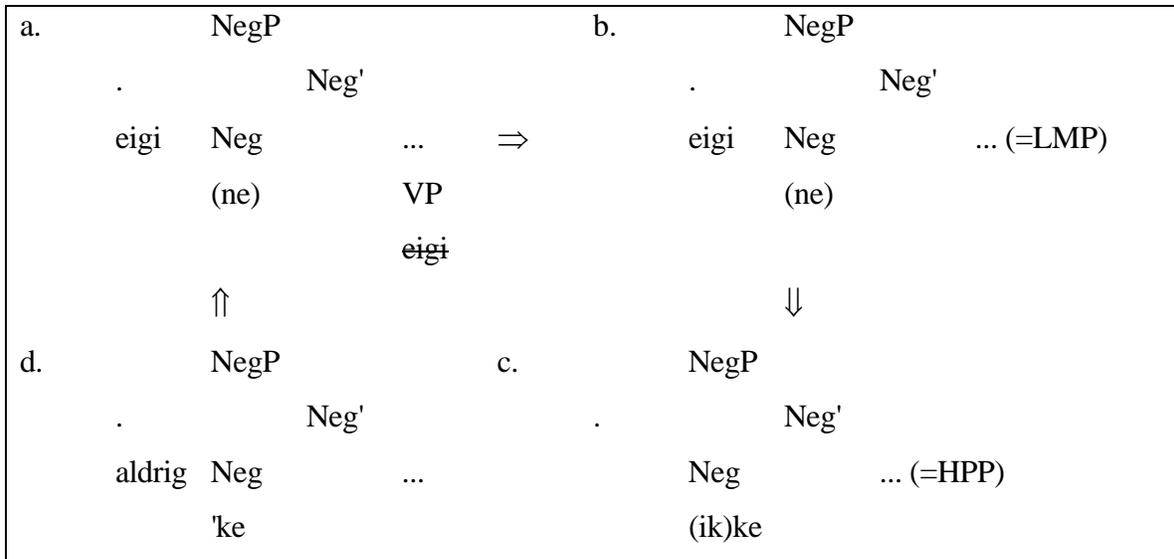


Figure 1: The Negative Cycle

Other examples of changes predicted by the HPP are given in (12):

- | | |
|---|---------------------------|
| (12) relative pronoun <i>that</i> to complementizer | Demonstrative to article |
| Negative adverb to negation marker | Adverb to aspect marker |
| Adverb to complementizer (e.g. <i>till</i>) | Full pronoun to agreement |

Under a Minimalist view of change, syntax is inert and doesn't change; it is the lexical items that are reanalyzed. Pronouns are reanalyzed from emphatic full phrases to clitic pronouns to agreement markers, and negatives from full DPs to negative adverb phrases to heads. This change is, however, slow since a child learning the language will continue to have input of, for instance, a pronoun as both a phrase and a head. Lightfoot (1999) develops an approach as to how much input a child needs before it resets a parameter. In the case of pronouns changing to agreement markers, there will have to be a large input of structures that provide evidence to the child that the full phrase is no longer analyzed as that. This is already the case in French, where in spoken French, the pronoun is always adjacent to the Verb. The child, therefore, always produces the pronoun in that position, even though regular subjects can precede or follow the verb (see Pierce 1992). However, the exact nature of the input needed for the change, the 'cue', is not explored in this paper.

Within Minimalism, there is a second economy principle that is relevant to grammaticalization. Combining lexical items to construct a sentence, i.e. 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). Initially, a distinction was made between merge and move and it was less economical to merge early and then move than to wait as long as possible before merging. This could be formulated as in (13):

(13) **Late Merge Principle (LMP):**

Merge as late as possible

In later Minimalism, merge is reformulated as external merge and move as internal merge, with no distinction in status. One could argue that (13) is still valid since the special Merge, i.e. internal Merge, requires steps additional to the ones Merge, i.e. external Merge, requires. The extra step is the inclusion in the numeration of copies in the case of internal Merge, e.g. a copy of *they* in (2). Traces are no longer allowed, since they would introduce new material into the derivation after the initial selection, and therefore copies of elements to be moved have to be included in the lexical selection. Move/internal merge is not just Move but `Copy, Merge, and Delete'. Since the numeration has to contain more copies of the lexical item to be internally merged, and since those copies have to be deleted in the case of traditional Move, (13) 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 that is expressed in the LMP.

The Late Merge Principle works most clearly in the case of heads. Thus, under Late Merge, the preferred structure would be (a) with *to* base generated in a higher position (here C but nothing hinges on that), rather than (b) with *to* in a lower position (here T) and moving to the higher position. See also Kayne (1999):

- (14) a. CP TP
 C . T'
 to . T
- b. CP TP
 C ^ . T'
 ... T ...
 to

This is indeed what has happened in a number of changes. There is evidence that Modern English speakers prefer (a) over (b):

- (15) a. It would be unrealistic **to not** expect to pay higher royalties (BNC-CSS 245).
 b. It would be unrealistic **not to** show them to be human (BNC-CBF 14312).

Corpus data show this preference, e.g. the adverb *probably* splits the infinitive in 22.7% of the cases in the (mainly written) British National Corpus, but in an American spoken corpus (CSE) it does so a 100%. The prescriptive rule against split infinitives is thus alive and most obvious in British written varieties. Such external rules interfere with Economy.

The preposition *for* underwent the same change from preposition indicating space to cause to non-finite complementizer. Examples are provided in (16) to (19):

- (16) *ouþer for untrumnisse ouþer for lauerdes neode ouþer for haeleste 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' (PC, anno 675).

- (17) *ac for þæm þe hie us near sint, we ... ne magon ...*

but for that that they us close are, we ... not may ...

`but because they are near to us, we can't ...' (*Orosius*, Bately 122.18-9).

- (18) *for æuric man sone ræuede oþer þe mihte*

because every man soon robbed another that could

`becasue everyone that could robbed someone else' (PC, 1135, 8).

- (19) *for agenes him risen sona þa rice men*

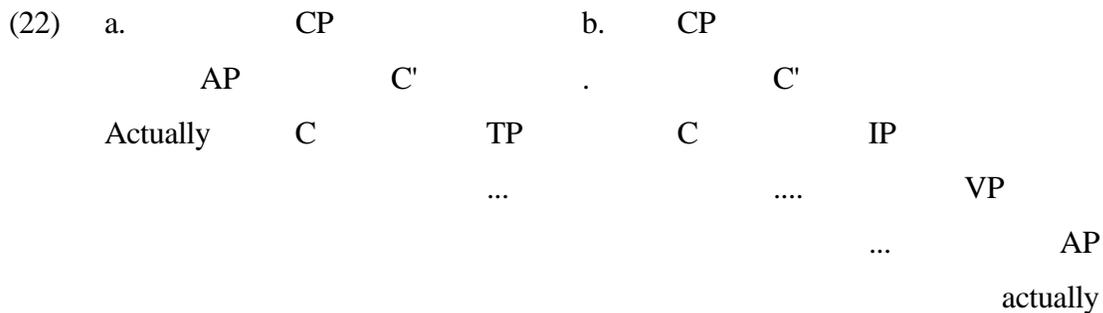
`because against him soon rose the powerful men' (PC, 1135, 18).

This 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). Late Merge also accounts for lexical **phrases** becoming base generated in the functional domain. An example is *actually*. When it is first introduced into the English language from French, it is as adjective (in 1315), and is then used as a VP adverb in the 15th century, meaning 'with deeds in actual reality' as in (20). It then changes to a higher adverb, as in (21), in the 18th century:

(20) Those who offend **actually**, are most grievously punished (OED 1660 example).

(21) **Actually**, it is kind of an interesting problem (CSE-FAC97).

Structure (22a) shows the more recent structural representation and (22b) the earlier one. The preferred one under the LMP is (22a):



Other examples of the LMP are given in (23):

- | | |
|--|--------------------------------------|
| (23) <i>Like</i> , from P > C (<i>like I said</i>) | Negative objects to negative markers |
| Modals: v > ASP > T | <i>To</i> : P > ASP > M > C |

How exactly does Late Merge account for language change? 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. Like the Head Preference Principle in (2), Late Merge is argued to be a motivating force of linguistic change, accounting for the change from specifier to higher specifier and head to higher head. Roberts & Roussou (2003), Wu (2000), and Simpson & Wu (2002) also rely on some version of Late Merge.

Concluding section 2, under the LMP as under the HPP, syntax is inert; it is the lexical items that are reanalysed. In this section, I have examined the Economy of Merge. Two principles, the HPP and the LMP, provide an insight into what speakers do when they construct a sentence. In the next section, I will apply these to a scenario for language evolution.

3 Grammaticalization and Language Evolution

In section 1, I have provided some background on how Merge could have been the first step in creating syntax from a stage that consisted of either words or gestures (e.g. Corballis 2000), or as Traugott (2004: 134) puts it "an exaptation of thematic role structure". The current section provides a scenario for subsequent steps.

Once Merge applies, certain structural and thematic relationships crystalize, as in (4) above. Another head (the small *v*) can be merged to structures as in (4), to accommodate the agent or causer in its Specifier. The *vP* represents the thematic level, and one that adult native speakers employ when they speak or write in 'fragments', as in (24). Children reach this stage too, as (25) shows, though they understand grammatical categories before they produce them:

(24) Work in progress

(25) want cookie

The next evolutionary stage is highly speculative, but it could be that the lower head starts to move to a higher head to express different relationships, e.g. *clear* from 'The sky is clear' to 'That clears the sky'. This is when the LMP applies as well and

grammatical elements arise, just as they do in the history of English in (16) to (19) above. One feature of the lexical element is emphasized over others (hence the slight semantic loss).

This same development can be seen in child language where children first master prepositions before using complementizers. Josefsson (2000: 398) shows that Swedish "children first acquire the PP and then, directly after that the subordinate clause". She divides the acquisition into Stage I with no prepositions, Stage II with occasional prepositions, and Stage III with first prepositions and then complementizers. "[M]ost often, the children do not start using complementizers at all until they have reached a 90% us of prepositions" in obligatory contexts, as in (26) and (27). These sentences provide a good illustration since the preposition and the complementizer have the same form *som*:

(26) *precis som en kan/ som en kanin*

just like a rab/ like a rabbit

(27) *grisen, den som heter Ola*

pig that who is-called Ola

(Embla, 27 months, both from Josefsson 2000: 410)

In English, we have a similar preposition, namely *like*. In fact, *like* can be a verb as well as in (28), from Abe (data from Kuczaj in the CHILDES database). What is interesting for a comparison with the Swedish data is that, at e.g. 3 years and 7 months old, Abe produces *like* many times as a verb and a preposition, as (28) and (29) show, but not as a complementizer. Later, e.g. at 4 years and 10 months, he uses *like* as a complementizer in (30). So the data are comparable to the Swedish, and the LMP works in child language as well:

(28) **like** a cookie (Abe, 3.7)

(29) no the monster crashed the planes down **like** this **like** that (Abe, 3.7)

(30) Daddy # do you teach **like** you do [/] **like** how they do in your school? (Abe, 4.10)

Similar data exist for *for* in English.

In conclusion, in this section, I have given some evidence from language acquisition that the HPP and LMP are part of the internalized grammar. I am proposing that language evolution followed a similar path, but we of course lack empirical evidence.

4 Renewal and the cycle

If the initial evolutionary stage of syntax is one where pragmatic relations are important (e.g. Bickerton 1990), the emergence of Merge will have the effect of incorporating the pragmatic material into a syntactic structure. This also occurs in grammaticalization-it is actually the start of the whole process (see Hopper and Traugott's 1993 cline). The two principles used in section 2 (HPP and LMP) take lexical material that is already part of the structure and change the position of it. There are also a number of changes where a new element comes from outside of the sentence, e.g. a special pronoun being incorporated into the CP to indicate subordination, and an emphatic topic pronoun becoming the subject (in Spec TP). This can be expressed by means of a principle that incorporates (innovative) topics and adverbials in the syntactic tree:

(31) **Specifier Incorporation Principle (SIP)**

When possible, be a specifier rather than an adjunct.

Sometimes, these 'renewals' are innovations from inside the language, as in the case of the English negative DP *na wiht* 'no creature' to mark negation but other times, these renewals are borrowed through contact with other languages. One such possible case is the introduction into English of the *wh*-relative. In Old English, there are a number of relative strategies, but by Early Middle English, the complementizers *þat* and *þe* are typical. This is predicted under the HPP since those forms are heads (see van Gelderen 2004: 83-7). By later Middle English, this form is 'competing with the *wh*-pronoun still present in present-day English (be it mainly in written English). Mustanoja

of the higher ASP(ect)P. The SIP would enable the Specifier position to be filled from the outside.

Givón (1979) and others have talked about topics that are later reanalyzed as subjects, and call this a shift from the pragmatic to the syntactic. What this means is that speakers tend to use the Phrase Structure rules, rather than loosely adjoined structures. With (31) added, typical changes can therefore be seen as (34):

- (34) a. Head > higher Head > 0 (=LMP)
 b. Adjunct > Spec > Head > 0 (=SIP/LMP and HPP)
 Phrase >

The change in (a) is the one from lower head (either lexical or grammatical) to higher head, via LMP. The change in (b) shows that either an adjunct (via SIP) or a lower phrase (via LMP) can be reanalyzed as specifiers, after which the specifier is reanalyzed as head (via HPP).

In this section, I have suggested that the emergence of syntax could have followed the path that current grammaticalization also follows and one that children take as well. In particular, Merge brings with it, a set of relations and a set of Economy Principles, from which grammaticalization and language change follow. The economy principles I have been discussing in section 2 are of the non absolute kind: if there is evidence for a pronoun to be both a phrase and a head, the child/adult will analyze it initially as head unless there is also evidence in the grammar (e.g. from coordination) that pronouns also function as full DPs.

5 Conclusion and response to critical views

I have looked at two steps that are required in the evolution from pre-syntactic language to language as we currently know it. The one is Merge and the structural and thematic relations it entails to build a basic lexical layer (the VP). The other is Economy of Merge, the HPP and LMP, the principles that enable learners to choose between different analyses. These

two principles result in what is known as grammaticalization and build the non-lexical layers (the TP and CP). Lexical material is also incorporated into the syntax through a third principle, the SIP. This principle allows the speaker to creatively include new material, e.g. as negative reinforcement in special stylistic circumstances.

Newmeyer (2006) voices some criticism of the view that grammaticalization tells us anything about the origins of language. His main reasons for skepticism are the alleged uni-directionality of grammaticalization and the exclusively lexical status of categories that an initial stage would show under this scenario. Newmeyer notes that some grammaticalizations from noun/verb to affix can take as little as 1000 years, and wonders how there can be anything left to grammaticalize if this is the right scenario. The Specifier Incorporation Principle proposed in the previous section, however, provides an answer for what the source of the replenishments are, namely borrowings and creative inventions. The Economy Principles do not provide a reason why certain languages/societies are more conservative than others, e.g. why the split infinitive has encountered such opposition by prescriptivists, and has kept *to* from grammaticalizing more. The reasons for this are sociolinguistic. A similar point raised by Newmeyer is that there are some languages that are argued to have undergone very little grammaticalization, and that "grammaticalization per se cannot tell us very much about the origin and evolution of language" (2006: 1). This is a point that concerns the status of languages described in work by e.g. Gil, and on which I cannot comment here, but see *Linguistic Typology* 9.3 (2005) for a critique of Gil.

Secondly, Newmeyer mentions that "there is no reason to assume that the earliest humans could not express concepts like 'in' and 'past time'" (2006: 1). This argument only works if one assumes that all early words were arguments, and not adverbials. If early humans had words for 'under the trees' or 'into a cave', they could express location and direction, and then these could later grammaticalize into prepositions.

Abbreviations

BNC	British National Corpus, see references.
CSE	Corpus of Contemporary Professional American English, see references.
HPP	Head Preference Principle

LMP	Late Merge Principle
OED	Oxford English Dictionary
SIP	Specifier Incorporation Principle
UG	Universal Grammar

References

- Abraham, Werner 1993. "Grammatikalisierung und Reanalyse: einander ausschließende oder ergänzende Begriffe". *Folia Linguistica Historica*, 13: 7-26.
- Bickerton, Derek 1990. *Language and Species*. Chicago: University of Chicago Press.
- British National Corpus, BNC, <http://sara.natcorp.ox.ac.uk>.
- Chomsky, Noam 1986. *Knowledge of Language*. New York: Praeger.
- Chomsky, Noam 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Chomsky, Noam 2002. *On Nature and Language*. CUP.
- Chomsky, Noam 2004. "Beyond Explanatory Adequacy". In Adriana Belletti (ed.), *Structures and Beyond*, 104-131 OUP.
- Chomsky, Noam 2005. "Three factors in Language design". *Linguistic Inquiry* 36.1: 1-22.
- Chomsky, Noam 2006. "Approaching UG from below". ms.
- Corballis, Michael 2002. "Did Language evolve from Manual Gestures?". In Alison Wray (ed.), *The Transition to Language*, 163-180. OUP.
- Corpus of Contemporary Professional American English, CSE, <http://www.athel.com>.
- Faarlund, Jan Terje 2004. *The Syntax of Old Norse*. Oxford: OUP.
- Gelderen, Elly van 2004. *Grammaticalization as Economy*. Amsterdam: John Benjamins.
- Gil, David 2006. "Early Human Language was Isolating-Monocategorical-Associational". ms.
- Givón, Tom 1979. "From discourse to syntax". *Syntax & Semantics* 12, 81-112. New York: Academic Press.
- Hale, Ken & Jay Keyser 2002. *Prolegomenon to a Theory of Argument Structure*. MIT Press.
- Heine, Bernd & Tania Kuteva 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.

- Hopper, Mike & Elizabeth Traugott 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Jelinek, Eloise 1998. "Voice and transitivity as Functional Projections in Yaqui". In Miriam Butt et al (eds), *The Projection of Arguments*, 195-224. Stanford: CSLI.
- Jespersen, Otto. 1922. *Language*. London: Allen & Unwin.
- 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. Kluwer.
- Kayne, Richard 1999. "Prepositions as Attractors". *Probus* 11, 39-73.
- Lambrecht, Knut 1981. *Topic, Antitopic, and Verb Agreement in Non Standard French*. Amsterdam: John Benjamins.
- Lightfoot David 1999. *The development of Language*. Malden: Blackwell.
- Mustanoja, Tauno 1960. *A Middle English Syntax*. Helsinki.
- Newmeyer, Frederick 2000. "On the Reconstruction of 'Proto-World' Word Order". In Chris Knight et al (eds) *The Evolutionary Emergence of Language*, 372-388. CUP.
- Newmeyer, Frederick 2006. "What can Grammaticalization tell us about the Origins of Language?". Abstract, <http://www.tech.plym.ac.uk/socce/evolang6/newmeyer.doc>
- Pierce, Amy 1992. *Language Acquisition and Syntactic Theory*. Dordrecht: Kluwer.
- Roberts, Ian & Anna Roussou 2003. *Syntactic Change*. Cambridge: Cambridge University Press.
- Romaine, Suzanne 1982. *Socio-historical Linguistics*. Cambridge: Cambridge University Press.
- Rydén, Mats 1983. "The Emergence of *who* as relativizer". *Studia Linguistica* 37: 126-134.
- Simpson, Andrew & Xiu-Zhi Zoe Wu 2002a. "Agreement Shells and Focus". *Language* 78.2: 287-313.
- Simpson, Andrew & Xiu-Zhi Zoe Wu 2002b. "From D to T - Determiner Incorporation and the creation of tense". *Journal of East Asian Linguistics* 11: 169-202.
- Traugott, Elizabeth 2004. "Exaptation and Grammaticalization". Minoji Akimoto (ed.) *Linguistic Studies based on Corpora*, 133-56. Tokyo: Hituzi Syobo.

Wu, Zoe 2004. *Grammaticalization and Language Change in Chinese*. London:
RoutledgeCurzon.

Notes

<1> Chomsky entertains both the possibility that syntax was "inserted into already existing external systems", namely the sensory-motor system and system of thought (Chomsky 2002: 108), as well as the one where the externalization develops after merge Chomsky (2006: 9-10).