The Relative Cycle in Hungarian Declaratives*

Julia Bacskai-Atkari

1 Introduction

The aim of this paper is to investigate the diachronic development of the Hungarian complementiser *hogy* 'that' and to account for its contribution to word order changes in the left periphery of various subordinate clauses.

In Modern Hungarian (from the end of the 18th century), *hogy* typically introduces finite complement clauses and is located in the higher C head position (cf. Rizzi 1997). Furthermore, it also appears as part of certain complex complementisers, such as *merthogy* 'because that'. By contrast, in Old Hungarian (9th century–early 16th century) *hogy* had various other functions as well, e.g. *hogy* was the original comparative complementiser. The reverse order of present-day complex complementisers also existed, e.g. *hogymert* 'that because'; also, *hogy* could potentially appear in relative clauses, together with relative operators (e.g. *hogy ki* 'that who').

As far as the meaning of these combinations is concerned, it must be noted that it was transparent in most cases: *hogy* did not contribute to the meaning of the ultimate combination, i.e. for any complementiser combination *hogy*+X or X+*hogy*, the meaning was – originally – invariably that of 'X'.

The word order variation with respect to complementisers is summarised in Table 1:

<insert Table 1 here>

* This research was funded by the German Research Fund (DFG), as part of the research project *The syntax of functional left peripheries and its relation to information structure*, and formerly also the Collaborative Research Centre *Information structure: The linguistic means for structuring utterances, sentences and texts* (SFB-632).

As can be seen, *hogy* could be combined with other complementisers, namely *ha* 'if', *mert* 'because' and *mint* 'as/than' and if a given combination existed in the XY order, then it seems to have invariably developed the mirroring YX order as well. In addition, the original meaning of such combination pairs was exactly the same.

One may therefore be tempted to say that there was free word order variation in Hungarian complementisers in that Old and Middle Hungarian not only allowed the co-presence of two C heads but the order of these heads was not fixed. By contrast, in Modern Hungarian there seems to be no such variation allowed: out of the combinations given in Table 1, only the ones marked in grey survive (i.e. *hogyha* 'that if', *merthogy* 'because that', *mintha* 'as if' and *minthogy* 'than that'). Essentially, such an analysis would allow the co-presence of two C heads in one left periphery both in Old/Middle and in Modern Hungarian and the only difference between the two stages would be that while the former arrangement involved interchangeable C heads, in the latter configuration the relative positions of the individual complementisers are fixed.

However, such a proposal would fail to account for the observation that though *hogy* is the last element in most combinations, it appears as the first one in *hogyha*, hence one would treat *hogy* as a higher C head in most cases but a lower one in this exceptional case. Furthermore, it should also be explained what constraint rules out free variation in Modern Hungarian but not in the previous periods.

To avoid having to encounter these theoretical problems, I will show that the chief difference between Old/Middle and Modern Hungarian lies in the fact that while the former allowed the co-presence of two C heads, the latter does not. In turn, I will show that this restriction derives from the grammaticalisation of all complementisers into higher C heads, hence there complementisers are no longer base-generated in the lower C position. In other words, Old/Middle Hungarian truly allowed two C heads in a single clause but the word order in these combinations was fixed – and in such a way that the relative position of the individual

complementisers is predictable from the different timing of their grammaticalisation into proper C heads.

However, with the upward movement of lower C heads to the higher C position the reverse surface order of the combinations also came into being, resulting in the apparent word order variation between C heads. These latter configurations ultimately grammaticalised into single (but morphologically complex) C heads and are still preserved in Modern Hungarian. Hence the apparent free word order variation is actually a result of fully predictable grammatical processes.

My proposal relies on two basic assumptions: on the one hand, I will show that *hogy* developed from an operator into a complementiser, which is in keeping with the general mechanism of the relative cycle (see, for instance, van Gelderen 2009). On the other hand, I will argue that *hogy* became a general marker of subordination in Old and Middle Hungarian. Both of these are of crucial importance in understanding word-order variation and change in the left periphery: depending on what stage of the relative cycle a given element – a complementiser or an operator – was in, it could occupy different positions in the left periphery. Hence both the absolute syntactic position of *hogy* and its relative position to other elements was subject to change. In addition, the fact that *hogy* appeared in a wide range of structures implies that it was likely to combine with several other complementisers too.

In the following, I will show that the word order variation (*hogy*+C vs. C+*hogy*) in Old and Middle Hungarian was mainly a result of the upward movement of the C heads other than *hogy*, and the fact that Modern Hungarian exhibits mostly C+*hogy* orders is a result of these combinations being fully grammaticalised. In other words, although the various orders at first sight may seem to be a result of free word order variation, it can be shown that both variation and change are fully predictable in terms of general economy principles. Though the present investigation will chiefly concentrate on Hungarian data, it has to be stressed that the

observations are relevant also for a more general understanding of complementiser ordering and combinations.

2 The data

First of all, let us consider the basic data in terms of functional and structural differences between Modern Hungarian and earlier periods. Since *hogy* 'that' was by far the most significant complementiser taking part in combinations, the issue of why it was easily combined with other elements must be addressed and the answer lies fundamentally in the functional flexibility of *hogy*.

Most functions of the complementiser *hogy* can be observed in Old and Modern Hungarian alike. First, the most basic function is that *hogy* introduces simple embedded declarative clauses, in examples like (1):

(1) a. Láttam, **(hogy)** esik az eső.
saw.1SG that rains the rain
'I saw it was raining.' (Modern Hungarian)

b. & felkèlè hog o̧hazaiaba mēnè o̧kèt
and up.rose.3SG that she.homeland.POSS.3SG.ILLATIVE go.COND.3SG she.two
menėvèl Moabitidiſnç videkebo̧l/
daughter-in-law.POSS.3SG.COM Moab.DAT country.POSS.3SG.ELATIVE
Mert hallotta vala **hog** vr tèkėntėttė volna
for heard.3SG was.3SG that Lord looked.3SG be.COND.3SG
o̧nėpėt & adot volna o̧nèkic ėtkèkèt
he.people.ACC and gave.3SG be.COND.3SG they.DAT dishes.ACC
'Then she arose with her daughters in law, that she might return from the country of Moab: for she had heard in the country of Moab how that the Lord had visited his people in giving them bread.' (Vienna Codex 1, 15th century)

As indicated, in this function the overt presence of *hogy* is optional: it alternates freely with a zero complementiser.

Second, *hogy* may also introduce embedded imperatives, as in (2):

(2) a. Azt mondták, **(hogy)** menjek Portóba.
that.ACC said.3PL that go.SUBJ.1SG Porto.ILLATIVE
'They told me to go to Porto.' (Modern Hungarian)

b. & kèʒdec kėrni **hog** èltauoʒnec o̧
and began.3PL ask.INF that off.leave.COND.3SG he
videkecbo̧l
country.POSS.3PL.ELATIVE
'And they began to pray him to depart out of their coasts.' (Munich Codex 40ra, 1466)

Again, *hogy* may freely alternate with a zero in this function.

Third, *hogy* is also the complementiser (optionally) introducing embedded *wh*-interrogatives, resulting in sequences of *hogy* + an interrogative pronoun:

(3) a. Azt kérdeztem, **(hogy)** mikor indulsz.
that.ACC asked.1SG that when leave.2SG
'I asked when you were to leave.' (Modern Hungarian)

b. valobiʒō o̧n ko̧ʒo̧tto̧c aʒ vtban vetǫko̧dtec vala
indeed self among.POSS.3PL the way.INESSIVE contested.3PL was.3SG
hog ki o̧ ko̧ʒo̧tto̧c nagob volna
that who he among.POSS.3PL greater be.COND.3SG
'for by the way they had disputed among themselves, who should be the greatest' (Munich Codex 45rb, 1466)

Fourth, *hogy* introduces clauses of purpose:

(4) a. Elmentem, **hogy** vegyek kenyeret.
off.went.1SG that buy.SBJV.1SG bread.ACC
'I went to buy some bread.' (Modern Hungarian)

b. & monda aʒocnac Meńńètec a rokon falucba / &
and said.3SG those.DAT go.SBJV.2PL the cognate villages.ILLATIVE and
varoſocba **hog** ot eſ p̄dical'l'ac mert arra
towns.ILLATIVE that there also preach.SBJV.1SG because that.SUBLATIVE
io̧ttèm
came.1SG
'And he said unto them, Let us go into the next towns, that I may preach there also: for therefore came I forth.' (Munich Codex 37ra, 1466)

Note that in this use, *hogy* cannot be replaced by a zero; the same is true for its fifth function, which is that of introducing resultatives: in these structures, the matrix clause contains a degree element (*úgy* 'so' or *olyan* 'so'), which selects for a subclause headed by *hogy*. This is illustrated in (5):

(5) a. Mari úgy elesett, **hogy** két hétig kórházban
Mary so off.fell.3SG that two week.TERMINATIVE hospital.INESSIVE
volt.
was.3SG
'Mary fell so badly that she spent two weeks in hospital.' (Modern Hungarian)

b. & ſokan gǫlekeʒen^{c} egbè / ug **hog** ſem a haʒba
and many gathered.3PL together so that neither the house.ILLATIVE
ſem aʒ aitohoz nē fėrnėn^{c}
neither the door.ALL not got.3PL
'And the multitude cometh together again, so that they could not so much as eat bread.' (Munich Codex 37rb, 1466)

In addition to these functions, which are shared between Old/Middle and Modern Hungarian, in Modern Hungarian there is a new environment, which is unattested in the previous periods: the introduction of embedded yes-no questions. In embedded yes-no questions, as illustrated in (6) below, the interrogative subclause invariably contains the question marker *-e* (usually attached to the verb), which is responsible for the marking of [+wh]; however, the overt complementiser *hogy* may also appear overtly, which shows that *hogy* is underspecified for the feature [±wh], see É. Kiss (2002: 99, 239). Consider:

(6) Azt kérdeztem, **(hogy)** éhes vagy-e.
that.ACC asked.1SG that hungry be.2SG-Q
'I asked whether you were hungry.' (Modern Hungarian)

On the other hand, there is an obsolete function of *hogy*, which appeared in earlier periods but not in Modern Hungarian: *hogy* was the complementiser introducing comparative subclauses, in subclauses expressing inequality followed by the element *nem* 'not', giving the sequence *hogy nem* 'that not':

(7) Mert iob hog megfog'doſuā algukmėg' vrat
because better that PARTICLE.catch.PTCP bless.SUBJ.1PL.PARTICLE Lord.ACC
hog **nė** mėg-hal'l'ōc
that not PARTICLE.die.SUBJ.1PL
'because it is better that we should bless the Lord when we are caught than to die' (Vienna Codex 25, 15th century)

This function was present both in Old and Middle Hungarian.

Apart from functional similarities and differences, the issue of combinations with other complementisers must also be addressed since this is the domain where complementiser word order variation may be observed. In Modern Hungarian, the following combinations are available: *minthogy* 'than that', *merthogy* 'because that' and *hogyha* 'that if' (note that here I intend to give meanings that reflect the morphological setup of the given combinations since

this will be crucially important for the present investigation). Historically, however, the reverse order of all of these combinations existed, alongside the ones mentioned above, hence: *hogymint* 'that than', *hogymert* 'that because' and *hahogy* 'if that'.

Before turning to the detailed analysis of how *hogy* took part in word order variations in the left periphery, let us see some preliminary data that show basic facts about the history of *hogy*. The research I carried out was based on a small corpus analysis whereby I examined the Gospel of Mark in three translations: the Munich Codex (1466) from the Old Hungarian period, György Káldi's translation (1626) from Middle Hungarian and the so-called Káldi-Neovulgata (1997), which is from Modern Hungarian. I examined altogether 230 loci: in all of these cases at least one of the translations contained the complementiser *hogy* or a combination thereof.[1] The most important results are summarised in Table 2:

<insert Table 2 here>

As can be seen, the number of *hogy* substantially increased over the periods in question; moreover, the number of zero alternates in these 230 loci also increased. Both follow from the fact that finite subordinate clauses became more frequent as opposed to non-finite clauses (cf. Haader 2001). Second, the combination *hogynem* 'that not' is present only in the Old Hungarian text but not in Middle or Modern Hungarian: this is in line with the fact that *hogy* was used as a comparative complementiser fundamentally in the Old Hungarian period and this function gradually came to be lost in later periods. Third, as can be seen, the combination *hogyha* 'that if' was already present in Old Hungarian and survives until the Modern

[1] Note that I found instances only of the combinations *hogynem* 'that not' and *hogyha* 'that if', hence there are no data for other combinations that were or are possible otherwise in a given period). I will return to the changes attested for complementiser combinations in section 5.

Hungarian period; hence at least some combinations of *hogy* originating in Old Hungarian are preserved in the language.

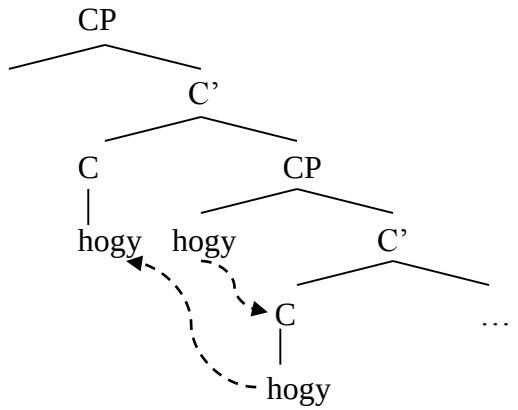
3 Grammaticalisation and the relative cycle

In order to provide an adequate syntactic analysis for the changes affecting *hogy* 'that', let us first discuss the notion of the relative cycle, as introduced by van Gelderen (2009), following the ideas of Hopper and Traugott (1993) or Heine and Kuteva (2002), among others. The relative cycle is a grammaticalisation process whereby a pronoun first becomes an operator moving to [Spec,CP] and subsequently this operator is reanalysed as the head of that CP.

In addition, there is a further possibility in terms of the grammaticalisation of C heads, which is the reanalysis from lower C to higher C. Both this process and the relative cycle are attested for English *that*, as discussed by van Gelderen (2009).

The same is true for Hungarian *hogy*; the processes are summarised in (8) below:

(8)



As can be seen, *hogy* was originally an operator moving to the lower [Spec,CP] position – at this stage, the meaning of *hogy* was 'how' (cf. Juhász 1991: 479–481, 1992: 781, 783–785, 801; Haader 1991: 729–737, 1995: 510–677).[2] In line with the general mechanism of the relative cycle, this operator was reanalysed as a complementiser, i.e. instead of an element moving to the C-domain it was reinterpreted as an element base-generated there, hence as a

[2] Evidence for this comes from early Old Hungarian data, see section 4.

(lower) C head.[3] Second, from a lower C head it was reinterpreted as a higher one, which is responsible for marking the Force of the clause (cf. Rizzi 1997).[4] In order for operators to develop into complementisers they have to have features that are compatible with C heads; in Hungarian, for instance, C heads are not allowed to have person and number features and hence ordinary relative operators (e.g. *aki* 'who') cannot be reanalysed as C heads, while operators such as *hogy* could since they had essentially the same features as complementisers. Note that both steps of reanalysis are motivated by economy, which can be summarised in the form of two principles, as described by van Gelderen (2004): the Head Preference Principle (HPP) and the Late Merge Principle (LMP), both originally going back to the idea that Merge is preferred over Move (cf. Chomsky 1995). The HPP states that being a head is preferable to being a phrase – hence the reanalysis from operator to complementiser. The LMP states that it is more economical to be base-generated in a higher position than to be moved to that position – hence the reinterpretation of the original lower C as a higher one.

The reason behind this latter step is simply that it is the higher C head that is responsible for defining the Force of the clause and the fact that certain overt lower C heads become associated with carrying Force implies that these elements also start moving up to the higher

[3] It is worth mentioning that reanalysis does not necessarily mean the complete loss of a previous function: it is possible to have a split between the original function and the newer one. This is also the case for *hogy*, which is preserved as an interrogative operator 'how' even in Modern Hungarian (with the possible alternate *hogyan*); the relative operator, however, developed a distinct form with *a-* during Late Old Hungarian and Early Middle Hungarian, hence Modern Hungarian has the forms *ahogy* (and *ahogyan*).

[4] Note that while the analysis of Rizzi (1997) suggests that there is a clear-cut difference between the higher and the lower C (Force and Fin, respectively), it is evident from Rizzi's work that this is not entirely the case: in fact, all Force heads identified by Rizzi (1997) are unambiguously associated with [+finiteness]. Furthermore, it is also possible to mark clause type lower than the highest CP node, as described by Rizzi (1999) for interrogatives. Given this, I will refer to the two CPs as higher and lower CP rather than Force and Fin.

C head. This again leads to a choice between movement and base-generation at a higher point in the structure – and just as in the case of the HPP, the latter configuration is preferred. Note that the preference of Merge over Move follows from the general principles of language acquisition, as described by Roberts and Roussou (2003: 202–218): the structure involving Merge is simpler both in derivational and in representational terms, since the element in question does not demonstrate feature syncretism. Therefore, when an element X can potentially be analysed as merged at a given point, or as merged at a lower level and moved subsequently higher up, the former option will be preferred by the language learner, that being more transparent. In this sense, the HPP and the LMP are merely epiphenomenal results of the actual driving forces underlying grammaticalisation processes; still, since there are some typical paths that grammaticalisation processes follow (as described by Roberts and Roussou 2003 in detail), the HPP and the LMP are useful notions for the description of the various changes that can be observed in the history of a language.

4 Simplex complementisers

Before turning to the actual complementiser combinations and word order variations in the left periphery, let us first briefly discuss the changes affecting other complementisers, which is necessary to understand how these could be combined with *hogy* 'that'.

In addition to *hogy*, Hungarian has three other complementisers that also developed by way of the relative cycle; these are *ha* 'if', *mint* 'than' and *mert* 'because'. All of them were originally operators (cf. Juhász 1991: 479–481, 1992: 781, 783–785, 801; Haader 1991: 729–737, 1995: 510–677). The operator function of *hogy* from Early Old Hungarian is illustrated in (9):

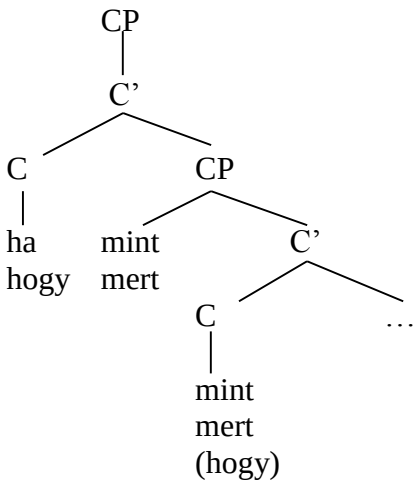
(9) furiſcte muſia!|| etetý ýmletí. ug **hug** ana ſciluttet.
bathes washes feeds breastfeeds so how mother child.POSS.ACC
'she bathes, washes, feeds and breastfeeds him as a mother does her child' (Königsberg Fragment, 14th century)

However, it is important to stress that there are crucial chronological differences between the individual operators. The functional split between operator and complementiser for *hogy* and

ha took place mostly in Proto-Hungarian: there are very few instances of *ha* used as an operator (and this function gradually disappeared altogether from the language), and while the operator *hogy* is attested in Old Hungarian too (and even later), the complementiser use can be detected already in the early documents. Furthermore, there is also reason to believe that they were already higher C heads as complementisers (though *hogy* could rarely be a lower one as well): this is shown by their relative positions in complementiser combinations (they come first in base-generated orders), as will be shown in the next section, and they also precede relative operators (see section 8); in addition, *hogy* also preceded the negative polarity head (located in between the two C heads) in comparative subclauses (see Bacskai-Atkari 2014: 213–218). By contrast, for *mint* and *mert* the split took place only during the Old and Middle Hungarian periods and hence these were either operators in the lower [Spec,CP] or were already base-generated in the lower C head as complementisers.

The possible Old Hungarian positions for present-day complementisers are summarised in (10) below (note that *ha* and *hogy* had already split from the original operator uses by early Old Hungarian, hence these operators uses are not included here):

(10)



As can be seen, the various present-day complementisers could take various positions historically; ultimately all of them came to be base-generated in the higher C head position.

5 Multiple complementisers

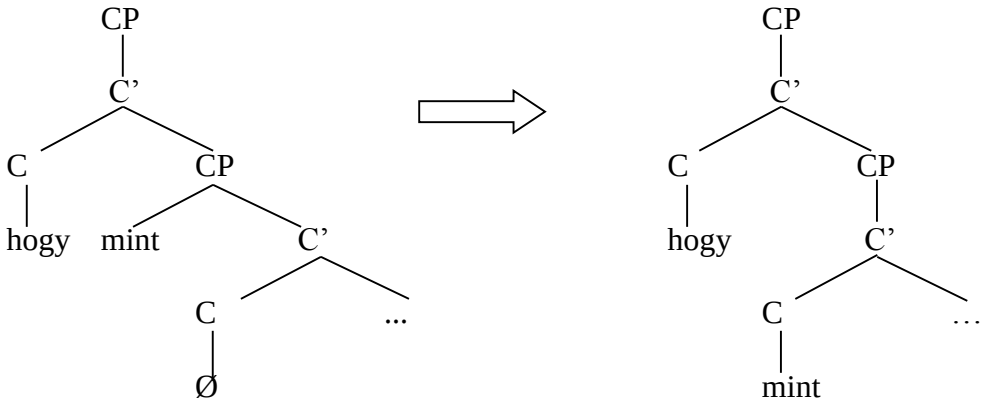
Due to the difference in the positions of complementisers mentioned above, the prediction is that it should be possible to have two overt complementisers – in predictable order – within one left periphery as long as the positional differences still held, i.e. in Old and Middle Hungarian. This prediction is borne out: it was possible to have a higher C head combined with a lower C head, ultimately deriving from original combinations of a higher C head and an operator located in the lower [Spec,CP] position.

The example in (11) shows the combination of *hogy* 'that' with *mint* 'than':

(11) edesseget erze nagÿoban **hogÿmint** annak elǫtte
sweetness.ACC felt.3SG greater that.than that.DAT before.POSS.3.SG
'(s)he felt sweetness even more than before' (Lázár Codex 140, after 1525)

The possible structures for such combinations are shown in (12):

(12)



The left-hand side diagram shows the configuration where the higher C head is headed by an overt complementiser (here: *hogy*) and the specifier of the lower CP is filled by an operator (here: *mint*), while the head of the lower CP is not filled by overt material. The operator was later reanalysed as the head of the lower CP, as shown in the right-hand side diagram. Hence examples such as (11) have either of these structures underlyingly.

While it is often difficult to decide whether a particular example shows an operator or a complementiser use of *mint* or *mert*, there are clear indicators that the operator use was lost towards the end of the Old Hungarian period, as *mint* and *mert* started to behave differently

from ordinary relative operators (such as *ki* 'who'). On the one hand, complementisers took part in head movement (see section 6), while relative operators did not. On the other hand, ordinary relative operators developed distinct forms from their interrogative operator counterparts, all relative operators starting with *a-* (hence *aki* 'who-Rel.' versus *ki* 'who-Int.'); however, this did not affect *mint* and *mert*, which indicates they were categorically distinct from ordinary relative operators by the end of the Old Hungarian period

Given that all complementisers had their respective typical positions in the left periphery, the expectation is that this should have a bearing on the possible word orders. This is indeed so: the word orders of C+C combinations – as well as combinations of a higher C head with an operator – were fixed. Since *hogy* was typically in the higher C head, it normally appeared as the first element in C+C sequences, hence the combinations *hogymint* 'that than' and *hogymert* 'that because'.

On the other hand, *ha* 'if' was invariably a higher C head; consequently, in combinations with *hogy*, *ha* was the first element, hence *hahogy* 'if that'. I will return to this issue in section 7; for the time being, suffice it to say that the word order of complementisers in combinations is fully predictable on the basis of their typical positions in the left periphery.

6 Complex complementisers

The question arises how the reverse word orders can be accounted for, given that the complementiser word order in C+C combination was fixed. As has been said, lower C heads were ultimately all reanalysed as higher C heads: this naturally involved the upward movement of these original lower C heads. Movement was motivated by the preference for marking the Force of the clause on the higher C head: however, complementisers first had to be reanalysed from operators, which are (and were) always located in the lower [Spec,CP] in Hungarian, hence complementisers had to appear in the lower C head first. While operators may indeed be the overt markers of clause type, they are not responsible for defining the Force of the clause, and hence are not necessarily associated with the higher CP cross-

linguistically: as far as Hungarian is concerned, they are in the lower CP in all periods, and they do not take part in upward movement, unlike lower C heads, which were preferably interpreted as not only overt markers of clause type but also as Force heads in language acquisition. The fixed order of original C+C combinations (as described in section 5) hence stems chiefly from the fact that certain complementisers grammaticalised later than others, which is attested independently from combinations, too.

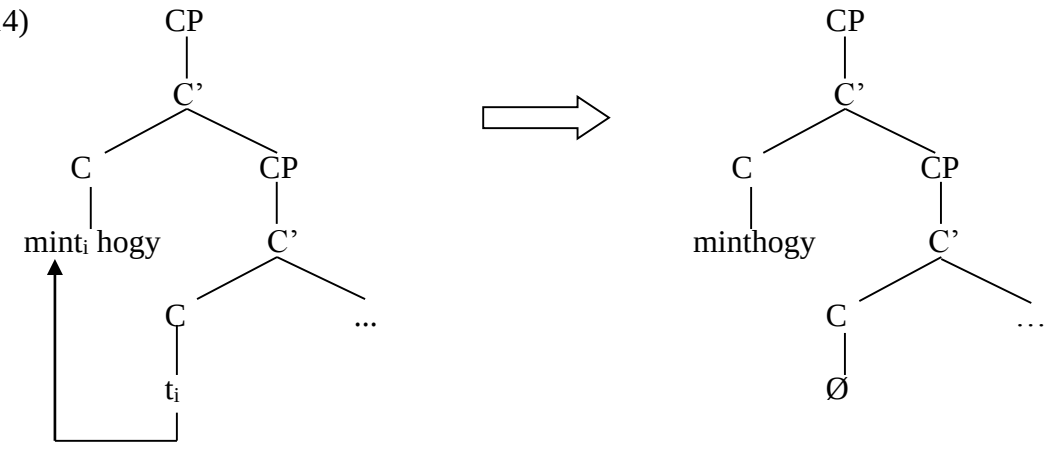
The upward movement of lower C heads was possible also when the higher C head was filled by another overt complementiser. If so, the lower head joined the higher one via adjunction: however, adjunction took place so that it produced the reverse linear order of the two heads, due the Linear Correspondence Axiom (Kayne 1994; but cf. also the Mirror Principle of Baker 1985, 1988). Hence in the case of an original C+C combination such as *hogymint* 'than that', the complex C head is of the form *minthogy* 'than that':

(13)	semi	nagob	nem	mondathatik:	**mint**	**hogh**	legọn	istenek
	nothing	greater	not	say.PASS.COND.3SG	than	that	be.SBJV.3SG	God.DAT
	ania							
	mother.POSS							

'nothing can be said to be greater than that she be the mother of God' (Tihanyi Codex 143, 1532)

The relevant structures of complex complementisers are shown in (14):

(14)



The left-hand side diagram shows the actual derivation of complex complementisers from multiple complementisers: the lower C head moves to the higher one (just as when it is a single overt C head) and adjoins to it from the left, resulting in the reverse of the original

linear word order. The right-hand side diagram shows the result of grammaticalisation: complex complementisers are base-generated as a single – though morphologically complex – unit, which produces a derivationally simpler structure than the one involving movement. Since, as has been argued for, the word order of C+C combinations was fixed, it is predictable that the order of C heads in a complex C is fixed as well; hence there are combinations such as *minthogy* 'than that' and *merthogy* 'because that', as well as *hogyha* 'that if' (note that I will return to the particular issue of *hogyha* in the next section).

The variation in the linear word order between C+C combinations and complex C heads, however, is no longer attested in the language. The reason behind this is that Modern Hungarian no longer has the possibility of accommodating two separate C heads in one left periphery: this is because all complementisers have been reanalysed as higher C heads and consequently there is no overt complementiser to be base-generated in the lower C position. Naturally, this should rule out the original C+C combinations, which is indeed the case: these are all extinct. On the other hand, the reverse order combinations are preserved: these are morphologically complex units base-generated as single heads.

7 The position of *hogy*

Let us now further investigate the role of *hogy* 'that' in word order variations in the left periphery. As should be obvious from the discussion so far, the underlying order was typically of the form *hogy*+X (X referring to a complementiser other than *hogy*); this was so because *hogy* was typically in the higher C head. Consequently, as all complex C heads represent the reverse order, generally it is the combinations of the form X+*hogy* that remained in the language up to Modern Hungarian.

The only exception to these generalisations is the case of *hogy* and *ha* 'if'. Here the underlying order was *ha+hogy*, as demonstrated by the fact that intervening elements could also potentially appear between the two heads:

(15) **Ha** késen **hogy** el nyugot az nap, hamar esọt váry
if late that off set.PST.3SG the sun soon rain.ACC expect.SBJV.2SG
'if the sun has set late, expect rain soon' (Cisio of Cluj-Napoca. G3, 1592)

As can be seen, there is an adverb (*késen* 'late') appearing between the complementisers *ha* and *hogy*, which are located in one and the same left periphery: it is still just one conditional clause that could not be potentially analysed as two separate clauses. It is true that Hungarian does not generally prefer elements (i.e. topics, foci) to appear as high as between the two C positions, and this seems to be valid for earlier periods as well (see É. Kiss 2014 on the evolution of the clausal left periphery). Still, the fact that it is possible to have an element there clearly indicates that *ha* and *hogy* must be in two distinct positions.

Hence in combinations with *ha*, *hogy* was originally base-generated as a lower C head, contrary to other combinations. The explanation behind this is simply that *ha* was always in the higher C head and hence *hogy* could only be base-generated in the lower C head, which was still an available option for *hogy* (see section 4). However, since typically *hogy* was also a higher C head already, movement typically took place, and thus the complex C head showing the reverse order (*hogyha* 'that if') was more frequent even in Old and Middle Hungarian than the base-generated order (*hahogy* 'if that').

Regarding the frequencies of the individual combinations discussed here, I conducted a corpus search based on the normalised part of the Old Hungarian Concordance corpus.[5] The results are summarised in Tables 3 and 4 below.[6] Table 3 shows the distribution of multiple complementisers involving *hogy*, see section 5:

[5] The corpus is available from http://omagyarkorpusz.nytud.hu/en-intro.html, where the metadata (date, token numbers) for the individual texts are also available. As of 28 December 2014, the normalised part of the corpus includes the texts given in Tables 3 and 4, and some shorter texts that contain no instances of complementiser combinations (many of these shorter texts do not contain complex sentences at all). Note also that the corpus is restricted to Old Hungarian only, and currently there is no searchable Middle Hungarian corpus at all.

[6] The token numbers always refer to the normalised version without punctuation marks.

<insert Table 3 here>

Table 4 shows the distribution of complex complementisers involving *hogy*, see section 6:

<insert Table 4 here>

The number of occurrences is higher for multiple complementisers in the case of *hogy* and *mint* only; even for *hogy* and *mert*, the complex complementiser structure is more frequent. This is in line with the hypothesis that once an original operator is reanalysed as a complementiser, it preferably moves up to a position that is associated with Force. The difference is most striking in the case of *hogy* and *ha*: there are no instances of *hahogy* in the normalised corpus, while *hogyha* is far more frequent than any other complementiser combination dealt with here, indicating that the frequency of the combination of *ha* and *hogy* in itself cannot be a reason for the lack of *hahogy*. However, if one takes into account that *hogy* was preferably a higher C head already (unlike *mint* and *mert*), it follows that the preference for the complex complementiser order was more significant than in the case of the other two combinations.

8 Relative clauses

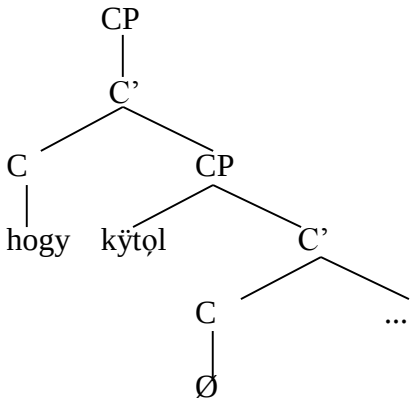
One of the most compelling questions in terms of the word order variations observed in connection with *hogy* 'that' is whether the analysis presented so far can be extended to other combinations involving *hogy* in the left periphery. A possible extension line is that of relative clauses, which could contain the sequence of *hogy* + a relative operator in Old and especially in Middle Hungarian (cf. Galambos 1907); however, the reverse word order (i.e. relative operator + *hogy*) is not attested. Hence the analysis should be able to account for the lack of surface word order variation here.

Consider the example for the combination *hogy* + operator in (16) below:

(16) olÿaat tezọk raÿtad **hog** **kÿtọl** felz
such.ACC do.1SG you.SUPERESSIVE that who.ABL fear.2SG
'I will do such a thing on you that you are afraid of' (Sándor Codex 28, 16th century)

As can be seen, the structure contains *hogy* besides the relevant form of the relative pronoun *ki* 'who'; the representation is given in (17):

(17)



The higher C head is filled by *hogy* and the specifier of the lower CP hosts the relative operator. Note that this configuration is identical to the left-hand side diagram in (12), that is, the configuration when a higher C head co-occurs with an operator that will later be reanalysed as a C head. However, in the case of relative operators such as *ki*, there is no such reanalysis: these operators never developed into complementisers since they never lost e.g. their person and number features that are incompatible with complementisers in Hungarian – on the other hand, the option was available for those operators that had essentially the same features as complementisers.

A further consequence of the lack of reanalysis in this case is of course that while sequences such as *hogy ki* 'that who' did occur, there are no instances of the reverse order, i.e. **ki hogy*

'who that': since these operators were not reanalysed as C heads, they did not (and could not) take part in head movement from lower C to higher C either.[7]

9 Further combinations

So far I have mainly been concerned with combinations involving two CP projections. Another point where the analysis can be extended is combinations that include negative-like Pol heads, which appear between the two CPs. The prediction is that if these combinations show surface word order variations (and changes), then these should be in line with the general mechanisms attested in original C+C combinations.

The base-generated C+Pol+C combinations were *hogynemmint* 'that not than' and *hogysemmint* 'that neither than'. Consider the following example containing *hogynemmint*:

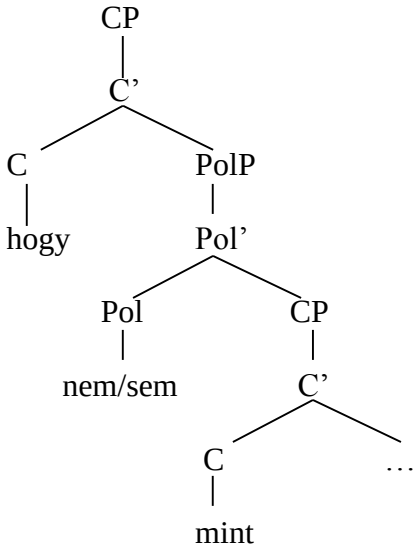
(18) az mentól alsobÿkban is tob angÿal uagon **honnem mÿnth** az
the more down.INESSIVE also more angel is that.not than the
napnak feneben
sun.DAT light.POSS.INESSIVE
'there are more angels in the basest one of them than in the sun's light' (SándK. 1v)

Assuming that the negative-like element *nem* is a Pol head that marks the negative polarity of the comparative subclause (cf. Bacskai-Atkari 2011; on the negative polarity of comparative subclauses cf. Seuren 1973; Gergel 2010; Matushansky 2011), the structure of the left periphery containing *hogynemmint* or *hogysemmint* should be the one given in (19) below:[8]

[7] Note that the motivation behind the upward movement of C heads is the marking of Force in the highest CP; since operators do not define Force, there is no motivation for an operator located in the lower [Spec,CP] to move further to the higher [Spec,CP] either.

[8] It has to be stressed that the negative-like element in comparatives like (18) is merely a polarity marker, and does not imply the presence of clausal (predicate) negation. Hence such combinations are different from Modern Hungarian *nehogy* 'lest', which contains a negative element (*ne*) and *hogy*: such purpose clauses are indeed instances of clausal negation. Note also that *nehogy* is not attested in Old Hungarian (see Haader 1995: 659; accordingly, a corpus search on the normalised part of the Old Hungarian Concordance corpus, as of 27 June 2015, did not yield any instances of *nehogy* either). It is highly unlikely that the combination involves the

(19)



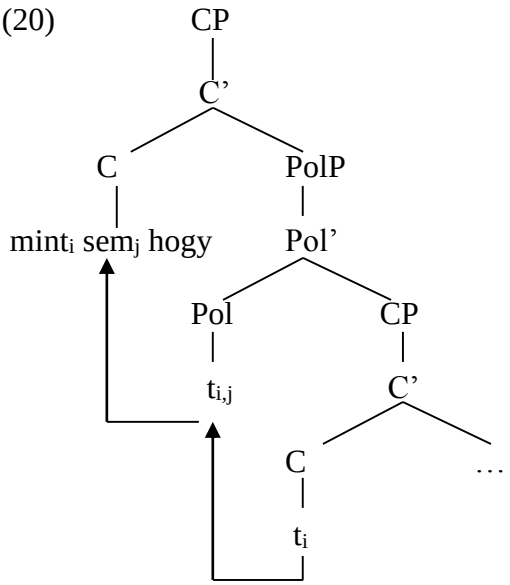
Hence there are two CPs and in between the two there is a PolP headed by the negative-like element *nem* ‘not’ or *sem* ‘neither’; while the higher C head is responsible for defining the Force of the clause, the overt, unambiguous marking of negative polarity and [+compr] is associated with lower functional heads. This configuration is no longer present in Modern Hungarian, which is predictable considering the assumption that complementisers are no longer base-generated in the lower C head. However, the reverse order combination survives in the form of *mintsemhogy* ‘than neither that’, which can be derived from the underlying order in the following way:

present-day declarative complementiser *hogy*, since *nehogy* in negative purpose clauses can be preceded by *hogy* (the presence of which is optional, just like in positive purpose clauses, see (4) in section 2):

(i) Elküldtem neked a kiállítás meghívóját, **(hogy) nehogy** lemaradj
off.sent.1SG you.DAT the exhibition invitation.POSS.ACC that lest miss.SBJV.2SG
róla.
it.DELATIVE
‘I sent you an invitation to the exhibition, lest you should miss it.’

Given these differences, I will not venture to provide an analysis for *nehogy* here, but since *nehogy* seems to be a single syntactic element below the C head expressing negation, it is highly unlikely that *nehogy* features a Pol+C combination.

(20)



The mechanism represented in (20) is in line with the proposal argued for in the present paper. That is, the lower C head *mint* ‘than’ first moves up to the intervening Pol head, filled by *sem* and adjoins to it from the left, resulting in the combination *mintsem* ‘than neither’. As a second step, the complex *mintsem* moves up to the higher C head since *mint* ultimately targets this position as the final landing site; the complex again adjoins to *hogy* ‘that’ from the left, giving the final combination *mintsemhogy*, which is thus the reverse order of the underlying configuration. Again, this complex head grammaticalised as such and hence is preserved even in Modern Hungarian.[9]

[9] Note that the same mechanism is not attested in the case of *nem*: this rather cliticised onto the preceding element *hogy*, and the resulting combination (*hogynem*, which was after phonological assimilation *honnem*) was probably interpreted as a complex head itself (for further details, cf. Bacskai-Atkari 2011, 2014: 213–218). While *hogy* was not an unambiguous overt marker for comparative Force, *hogynem* was, and it seems likely that there was a trigger for the upward movement of *mint* in the former case but not in the latter. Since the further discussion of this issue would require an investigation of comparatives, I will not venture to examine the question here any further.

10 Functions of *hogy*

Finally, let us briefly examine the changes in the functions of *hogy* 'that', as reflected by its role in complementiser combinations. It should be obvious that *hogy* was readily combined with several other elements (either C heads or operators) and thus played a crucial role in word order variations and changes in the left periphery. The question is what allowed for the high potential of *hogy* to appear in several types of subclauses and combinations.

Recall first the data from Table 2, which data show that the frequency of *hogy* gradually increased over the periods in question, and while certain combinations (e.g. *hogynem* 'that not') disappeared from the language, others (e.g. *hogyha* 'that if') survive until Modern Hungarian.

The major changes in the functions of *hogy* can be summarised as follows. Originally, in addition to introducing various embedded clauses that survive into Modern Hungarian, *hogy* had the specific function of introducing comparative subclauses, which was lost alongside the appearance of *mint* 'than/as'. On the other hand, *hogy* became the marker of subordination, which is also indicated by the fact that in Modern Hungarian its appearance is extended to embedded yes/no questions.[10]

The significance of marking subordination was in line with a general increase in the number of finite clauses. Table 2 showed that the number of clauses containing *hogy* – and its zero alternant – increased; I examined all the loci in question in all the three texts and looked at the various possible constructions that could stand instead of *hogy*-clauses in the Munich Codex and in Káldi's translation. The results are summarised in Table 5 below:

[10] Just as in the case of clause typing and finiteness, the marking of subordination also has a preferred relative position in the left periphery, which is the highest functional projection. The order of the three is identified as Sub > Force > Fin by Haegeman (2010). Since I do not wish to adopt a strict cartographic approach, I do not label CPs as any of these more specific categories, but the analysis presented here is compatible with the general theory.

<insert Table 5 here>

As can be seen, in Old Hungarian the significance of non-finite structures, as well as coordination and phrasal – i.e. mostly nominal – equivalents of finite subclauses was high and it decreased in the Middle Hungarian period (cf. Haader 2001). The possibility of having *mert* 'because' as a simple declarative subordinator was absent already from the Middle Hungarian period (cf. Haader 2003: 506).

The hypothesis that *hogy* became a general marker of subordination is supported by two phenomena. First, it appeared in a wide range of clauses, i.e. conditionals, clauses of reason, relative clauses, hence it does not seem to have been restricted to specific functions. Second, the meaning of a combination *hogy*+X or X+*hogy* did not initially differ from the meaning of X on its own (X standing for any complementiser different from *hogy* and also for ordinary relative operators), which indicates that *hogy* did not induce any semantic change by its appearance and thus its function was merely that of marking subordination. The fact that *hogy* was fundamentally a subordination marker also explains why it was able to appear in a wide range of structures and, consequently, why it frequently took part in surface word order variations in the left periphery.

11 Conclusion

The aim of this paper was to provide an analysis for the diachronic development of the Hungarian complementiser *hogy* 'that' and to show that the word order variations and changes attested in connection with it can be explained by considering the general grammaticalisation processes responsible for the changes affecting all complementisers in Old and Middle Hungarian. As was shown, *hogy* was reanalysed from an operator into a C head, the process of which can be traced in the case of other Hungarian C heads too.

Regarding the combinations of various complementisers, it was seen that the order in base-generated combinations follows from the differences in when each C head grammaticalised from an operator. In turn, the reverse order combinations can all be derived from the base-generated ones via head adjunction, which predicts that the linear order will be reversed. These complex complementisers fully grammaticalised and they remain in the language even in Modern Hungarian, while C+C combinations disappeared because complementisers are no longer base-generated in the lower C head.

Though all of these changes affect the entire complementiser system as such, the importance of *hogy* in this respect must be stressed since most combinations involved this particular element: this, as was seen, is because *hogy* was a general subordinator head that thus had the possibility to appear in a wide range of structures and to combine with other elements.

List of textual sources

Birk Codex. 1474.

Bod Codex. First half of the 16th century.

Booklet on the dignity of the apostles. 1521.

Cisio. Cluj-Napoca 1592.

Czech Codex. 1513.

Festetics Codex. 1492–1494.

Guary Codex. Before 1508.

Jókai Codex. Around 1440.

Kazinczy Codex. 1526–1541.

Königsberg Fragment and its Ribbons. Middle or second half of the 14th century.

Lázár Codex. After 1525.

Miskolc Fragment. 1525.

Munich Codex. 1466.

Praise of Saint Nicholas. Written by Ferenc Péchy. 1529.

Sándor Codex. The first quarter of the 16th century.

Song on the Conquest of Pannonia. Written by Demeter Csáti. Around 1526.

Székelyudvarhely Codex. 1526–1528.

Tihanyi Codex. 1532.

Vienna Codex. Middle of the 15th century.

References

Bacskai-Atkari, Julia (2011). 'A komparatív operátor esete a mondatbevezetővel: Szintaktikai változások a magyar hasonlító mellékmondatokban' [The comparative operator's affair with the complementiser: Syntactic changes in Hungarian comparative subclauses], in Katalin É. Kiss and Attila Hegedűs (eds.), *Nyelvelmélet és diakrónia*. [Language theory and diachrony] Budapest & Piliscsaba: Szent István Társulat, 103–119.

Bacskai-Atkari, Julia (2014). The Syntax of Comparative Constructions: Operators, Ellipsis Phenomena and Functional Left Peripheries. Potsdam: Universitätsverlag Potsdam.

Baker, Mark (1985). 'Mirror Theory and Morphosyntactic Explanation', *Linguistic Inquiry* 16.3: 373–415.

Baker, Mark (1988). *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.

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

É. Kiss, Katalin (2002). *The Syntax of Hungarian*. Cambridge: Cambridge University Press.

É. Kiss, Katalin (2014). 'The Evolution of Functional Left Peripheries in the Hungarian Sentence', in Katalin É. Kiss (ed.), *The Evolution of Functional Left Peripheries in Hungarian Syntax*. Oxford: Oxford University Press, 9–55.

Galambos, Dezső (1907). Tanulmányok a magyar relatívum mondattanáról. [Studies on the syntax of Hungarian relatives] Budapest: Athenaeum.

Gelderen, Elly van (2004). *Grammaticalization as Economy*. Amsterdam: John Benjamins.

Gelderen, Elly van (2009). ‘Renewal in the Left Periphery: Economy and the Complementiser Layer’, *Transactions of the Philological Society* 107.2: 131–195.

Gergel, Remus (2010). ‘Towards Notions of Comparative Continuity in English and French’, in Anne Breitbarth et al. (eds.), *Continuity and Change in Grammar*. Amsterdam: John Benjamins, 119–144.

Haader, Lea (1991). ‘Az alárendelő mondatok: Az alanyi, állítmányi, tárgyi és határozói mellékmondatok’ [Subordinate clauses: Subjective, predicative, objective and adverbial subclauses], in Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana I.: A korai ómagyar kor és előzményei*. [The diachronic grammar of the Hungarian language I: The Early Old Hungarian period and its antecedents] Budapest: Akadémiai Kiadó, 728–741.

Haader, Lea (1995). ‘Az alárendelő mondatok: Az alanyi, állítmányi, tárgyi és határozói mellékmondatok’ [Subordinate clauses: Subjective, predicative, objective and adverbial subclauses], in Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana II/2.: A kései ómagyar kor: Mondattan. Szöveggrammatika.* [The diachronic grammar of the Hungarian language II/2: The Late Old Hungarian period: Syntax. Textual grammar] Budapest: Akadémiai Kiadó, 506–665.

Haader Lea (2001). ‘Mikrodiakrónia és változásvizsgálat (az összetett mondatokban)’ [Microdiachrony and the analysis of changes (in complex sentences)], *Magyar Nyelvőr* 125: 354–371.

Haader, Lea (2003). ‘Az ómagyar kor: Mondattörténet: Az összetett mondat’ [The Old Hungarian period: The history of sentences: The complex sentence], in Jenő Kiss and Ferenc Pusztai (eds.), *Magyar nyelvtörténet*. [The history of the Hungarian language] Budapest: Osiris Kiadó, 500–560.

Haegeman, Liliane (2010). ‘Evidential Mood, Restructuring, and the Distribution of Functional *Sembrare*’, in Paola Benincà and Nicola Munaro (eds.), *Mapping the Left*

Periphery: The Cartography of Syntactic Structures, Volume 5. Oxford: Oxford University Press, 297–326.

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

Hopper, Mike and Elizabeth Traugott (1993). *Grammaticalization*. Cambridge: Cambridge University Press.

Juhász, Dezső (1991). 'A kötőszók' [Conjunctions], in Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana I.: A korai ómagyar kor és előzményei*. [The diachronic grammar of the Hungarian language I: The Early Old Hungarian period and its antecedents] Budapest: Akadémiai Kiadó, 476–500.

Juhász, Dezső (1992). 'A kötőszók' [Conjunctions], in Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana II/1.: A kései ómagyar kor: Morfematika*. [The diachronic grammar of the Hungarian language II/1: The Late Old Hungarian period: Morphology] Budapest: Akadémiai Kiadó, 772–814.

Matushansky, Ora (2011). *No more no less: existential comparatives revisited*. Ms. University of Utrecht.

Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT.

Rizzi, Luigi (1997). 'The Fine Structure of the Left Periphery', in Liliane Haegeman (ed.), *Elements of Grammar*. Dordrecht: Kluwer, 281–337.

Rizzi, Luigi (1999). *On the Position "Int(errogative)" in the Left Periphery of the Clause*. Available at: www.ciscl.unisi.it/doc/doc_pub/int.doc (Last accessed: 31 March 2008).

Roberts, Ian and Anna Roussou (2003). *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.

Seuren, Pieter (1973). 'The Comparative', in Ferenc Kiefer and Nicolas Ruwet (eds.), *Generative Grammar in Europe*. Dordrecht: D. Reidel, 528–564.