The History of Hungarian Complex Complementisers
Júlia Bácskai-Atkári
(Research Institute for Linguistics of the Hungarian Academy of Sciences)

Abstract

The article investigates the diachronic development of complementisers in Hungarian, with particular stress on how complex complementisers were combined from simplex ones. Using a minimalist framework, I show that the processes can be explained by the relative cycle, whereby original operators were reanalysed as complementiser heads, and were constrained by Kayne’s Linear Correspondence Axiom, which determined the order of two elements in head adjunction. The analysis explains why configurations having two separate C heads X and Y had to disappear from the language, while the ones derived via movement and showing a YX order, could remain.

1 The problem

In Modern Hungarian, there are four complementisers introducing finite subordinate clauses: ha ‘if’, hogy ‘that’, mert ‘because’ and mint ‘than/as’. These could combine historically in several ways, as shown in Table 1:

<table>
<thead>
<tr>
<th></th>
<th>ha</th>
<th>hogy</th>
<th>mert</th>
<th>mint</th>
</tr>
</thead>
<tbody>
<tr>
<td>ha</td>
<td>–</td>
<td>hahogy</td>
<td>–</td>
<td>hamint</td>
</tr>
<tr>
<td>hogy</td>
<td>hogyha</td>
<td>–</td>
<td>hogy mert</td>
<td>hogymint</td>
</tr>
<tr>
<td>mert</td>
<td>–</td>
<td>merthogy</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>mint</td>
<td>mintha</td>
<td>minthogy</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Table 1: The combinations of Hungarian complementisers

Considering the above data, there are several questions that emerge. First, certain theoretically possible combinations do not exist – naturally, an element does not combine with itself but there are still other logically possible configurations, such as e.g. mert + mint. Second, the pattern is completely symmetrical: if a combination is

1 The present research was funded by the project OTKA-78074. I would like to thank István Kenesei, Katalin É. Kiss and Gergely Kántor for their comments on the proposal. For his helpful remarks and questions at the conference, I am especially grateful to Markus Schiegg.
possible in the order XY, it is also possible in the order YX. Furthermore, it has to be mentioned that the original meaning of a combination XY is the same as that of YX. Third, though all of these combinations used to exist in the history of the Hungarian language, it is only the ones in the left bottom part of the chart (i.e. *hogyha, merthogy, mintha* and *minthogy*) that survive into Modern Hungarian: their counterparts showing the opposite order disappeared from the language. All of these regularities suggest that there must be rules responsible for these phenomena.

My proposal is that the four C heads were in different stages of development in Old Hungarian, which resulted in a fixed underlying order – conversely, the reverse order was made possible via movement.

To provide a theoretical framework for the investigation, section 2 will briefly describe the structure of the Left Periphery. Section 3 will be devoted to the etymology of the four simplex complementisers, which in section 4 will be followed by describing the historical development thereof. Sections 5 and 6 will deal with the evolution of complex complementisers; finally, I will summarize the proposed analysis in section 7.

2 The Structure of the Left Periphery

I will adopt the analysis of Rizzi (1997: 297; 2004: 237–238), according to which the structure of the Left Periphery is as follows:

![Figure 1: The structure of the Left Periphery](image)

As can be seen, there are two C heads, one responsible for Force and the other for Finiteness; in between the two, various Topic and Focus phrases may optionally occur, but these will not be relevant for our discussion now. Operators, as indicated, move to the specifier position of the lower CP via ordinary *wh*-movement (Chomsky 1977: 87; Kennedy and Merchant 2000: 89–90).

There are two important constraints to mention here in connection with this. First, in Modern Hungarian, the two C heads cannot be filled at the same time; in this respect, Hungarian is similar to Italian, as described by Rizzi (1997). That this is not necessarily so is shown by Welsh, which does allow two filled C heads:

(1) *Dywedais, i mai ‘r dynion fel arfer a werthith y ci.*

say I that the men as usual that sell the dog

‘I said that it’s the men who usually sell the dog.’

(ex. from Roberts 2005: 122)
This provides evidence for the two possible positions for C heads. Naturally, the question arises whether Old Hungarian resembled the parametric setting of Modern Hungarian and Italian, or rather that of Welsh.

Second, there is the Doubly Filled COMP Filter, which prohibits the co-presence of overt material in the specifier and an overt complementiser in the same CP.

I assume that in Modern Hungarian all the four complementisers are in the higher C head position, thus:

```
CP
  C'
    C_Force
    CP
        C'
          C_Fin
            ...
```

Figure 2: The position of complementisers in Modern Hungarian

One of the major questions is of course whether the position of complementisers did actually vary throughout the history of Hungarian, and if so, how. In order to provide an answer, let us first see the etymology of these complementisers.

### 3 The etymology of complementisers

Originally, present-day Hungarian complementisers were pronouns, which eventually came to be used as operators (Juhász 1991: 479–481, 1992: 781, 783–785, 801; Haader 1991: 729–737, 1995: 510–677). As a result of functional splits, these present-day complementisers still have etymologically related operators: hogy ‘that’ has hol ‘where’, ha ‘if’ has hová ‘where to’, mint ‘than/as’ has miként and miképpen ‘how’ and mert ‘because’ has miért ‘why’.

The splits, however, took place in different periods: while for hogy and ha, the split had already ended before the Old Hungarian period, in the case of mint and mert it happened during the Old Hungarian and the Middle Hungarian period. Consequently, for instance, a form miért could be used for both ‘why’ and ‘because’ and the same is true for mert, while in Modern Hungarian miért in invariably ‘why’ and mert is ‘because’. This consideration will be important for the analysis of their historical development.

Interestingly, though, new related operators started to appear in Old Hungarian for hogy and ha: these were hogy ‘how’ and ha ‘when-Rel.’, which are homonymous with their complementiser etymons – but their positions were different, as will be seen in the next section.

### 4 The history of simplex complementisers

The history of simplex complementisers involves the development from operators to C heads via reanalysis. As a second stage of reanalysis, these complementisers were reinterpreted from C_Fin heads into C_Force heads.
The process can be best described with the notion of the relative cycle. The relative cycle is a grammaticalization process, whereby an original determiner becomes first a relative operator, and subsequently the relative operator is reanalysed as a C head (Roberts–Roussou 2003: 119; van Gelderen 2009). In other words, an element moving to the CP domain is reinterpreted as one that is base-generated there – hence, as a head.

This kind of change happened to the English *that* during the Old English period: the element was originally a determiner (this function being preserved in the D head in Present-day English as well) but was used also as a relative pronoun. However, the relative pronoun moving to [Spec; CP] came to be analysed later as part of the CP, i.e. as a C head. First *that* was interpreted as a C\_Fin head and subsequently was reanalysed from the C\_Force head, as shown by van Gelderen (2009: 159).

The process described above is schematically drawn below:

![Diagram](image)

**Figure 3:** The development of *that*

As can be seen, the element *that* first occupied the specifier position of the lower CP, then it was reanalysed as the head thereof, and finally was base-generated as the head of the higher CP. Both steps are induced by economy, which can be described in terms of two major principles: the Head Preference Principle (HPP) and the Late Merge Principle (LMP), as described by van Gelderen (2009: 136; 2004). The first one says that being a head is preferable to being a phrase (which follows from a preference for merge over move, see also Chomsky 1995) – hence the change from an operator moving to [Spec; CP] to a C head. The latter establishes that merge (i.e. the insertion of new elements into the structure) should happen as late as possible – hence the preference for the C\_Force position over C\_Fin.

This is exactly what happened to the four Hungarian complementisers in question. However, there is an important chronological difference to be observed. In the case of *mint* ‘than/as’ and *mert* ‘because’, the operator function was still preserved in Old Hungarian, alongside the new one (i.e. that of C heads). By contrast, the old operator functions of *hogy* ‘that’ and *ha* ‘if’ were already lost in the period, and the existing operator functions were actually new: these are *hogy* ‘how’ and *ha* ‘when-Rel.’. The latter was rarer even during the Old Hungarian period but the former is still possible in Modern Hungarian:

(2) *Láttam, hogy úszik a dinnyehéj.*

saw-1.Sg. that/how drifts the melon rind

‘I saw that/how the melon rinds were drifting.’
The example in (2) has two readings precisely because in Modern Hungarian *hogy* ‘that’ and *hogy* ‘how’ are homonyms, which was the case in Old Hungarian too.

5 Two complementisers in one Left Periphery

The distinction between *hogy* ‘that’ / *ha* ‘if’ and *mint* ‘than’ / *mert* ‘because’ has an important consequence with respect to the positions these elements occupied. Since *hogy* and *ha* developed into C heads earlier, in Old Hungarian they were located in the upper C position as $C_{\text{Force}}$ heads. By contrast, *mint* and *mert* were later developments and therefore were either in the lower [Spec; CP] as operators or in $C_{\text{Fin}}$ as complementiser heads.

This resulted in a fixed underlying order of possible combinations of these elements, i.e. if there were two complementisers in one Left Periphery, the order obligatory had to be that of an upper C head followed by a lower one. Hence, this is the way how combinations like *hogy* *mint* ‘that than’, *hamint* ‘if as’, *hogy* *mert* ‘that because’ and *hahogy* ‘if that’. The last one, seemingly combining two upper C heads, is less straightforward to be mentioned in this category as the others: I will return to the question of why it still has to be here in section 7. For the time being, let us consider some examples for these combinations.

The combination *hogy* *mint* ‘that than’ was found in comparative subclauses (for a detailed analysis, see Bácskai-Atkári 2011):

(3) *edesseget erze nagyoban hogymint annak elotte*  
sweetness-Acc. felt-3.Sg. greater that.than that-Dat. before-Poss.1.Sg.  
‘(s)he felt sweetness even more than before’  
(LázK. 140; ex. from Haader 1995: 562)

(4) *hogy* *mint* akki zonetlen a kereztfanac o keférkeget v  
that as who constantly the cross-Dat. (s)he bitterness-Acc. (s)he  
telieben viželi  
body-Ine. bears  
‘as one who constantly bears the bitterness of the cross in his/her body’  
(NagyszK. 40–41; ex. from Haader 1995: 619)

As shown by (4), *hogy* *mint* could also appear in comparatives expressing equality, not just in ones expressing inequality, as in (3).

The string *hamint* ‘if as’ was used in conditional comparatives:

(5) *de hamjang* <ak el aluttak volna lelöketh istennek*  
but if as only PREV slept-3.Pl. be-Cond.3.Sg. soul-Poss.3.Pl.Acc. God-Dat.  
meg adaak  
PREV gave-3.Pl.  
‘but as if they had only fallen asleep, they gave their souls to God’  
(SándK. 28)

(6) *ha mynt az meennyey eedes elederre yarywlt vona ol*  
if as the heavenly sweet food-Subl. went.up-3.Sg.be-Cond.3.Sg. so  
eremelth  
willingly  
‘as if (s)he had so willingly gone there for the heavenly sweet food’  
(ÉrdyK. 282; ex. from Haader 1995: 637)
Turning now away from comparatives, *hogy mert* ‘that because’ appeared in clauses of reason:

(7) *Dehogý mert* zentferenc ỳgen zeretiual a ewtett týjtasagert es

but that because St Francis well liked was-3.Sg.him purity-Fin. and alazatossaargt

humidity-Poss.3.Sg.Fin.

‘but because Saint Francis liked him well for his purity and for his humility’

(JókK. 46)

(8) az en dolgom ebbe vagon *hogy merth* ferdinandof kýral

the I case-Poss.1.Sg. this-III. is that because Ferdinand king

folgalathomath megh nem ffýzethe

service-Poss.1.Sg.Acc. PREV not paid-3.Sg.

‘my case is that because king Ferdinand did not pay for my services’

(Mny. 2: 211; ex. from Haader 1995: 651–652)

Finally, the combination *hahogy* ‘if that’ was also made up of two separate complementisers and appeared in conditional clauses:

(9) *Az én jó istenem, ha hogy* sok ellenség, reám fegyverkezék,

the I good God-Poss.1.Sg. if that many enemy I-Subl. arm.

tolok megmente

they-Abl. saved-3.Sg.

‘my good God, if many enemies armed against me, saved me from them’

(B. Balassa)

(10) *Ha hogy* az ő keserves kin-szenvedését gyakor emlékezettel

if that the (s)he bitter torture-Poss.3.Sg.Acc. often memory-Com.

szivedben forgatod

heart-Ine. turn-2.Sg.

‘if you often remember his/her bitter torture in your heart’

(Csúzi:Síp. 105)

As has already been said, I will return to the issue of structures involving *hahogy*, such as (9) and (10) above. The combinations dealt with in this section were the ones that have two separate C heads at the beginning of a subclause – and also the ones that do not survive into Modern Hungarian.

The proposal so far predicts that combinations reflecting the underlying order of two C heads will exist. In such cases, these configurations could initially have the C_{Force} head followed by an operator and later by a C_{Fin} head, as the operators came to be reanalysed as such. The fact that C_{Force} heads combined with operators in this way is far from being unprecedented in the period: in Old Hungarian, and especially in Middle Hungarian, *hogy* and *ha* frequently combined with relative operators, resulting in strings like *hogy ki* ‘that who’, *ha ki* ‘if who’ or *ha mi* ‘if what’ (Juhász 1992: 792; Galambos 1907: 14–18; Bácskai-Atkári 2011: 112–113). Consider the following examples:

(11) *olýaat tezék raýtad hog kýtclé felz*

such-Acc. do-1.Sg. you-Sup. that what-Abl. fear-2.Sg.

‘I will do such on you that you fear’

(SándK. 14v)
The History of Hungarian Complex Complementisers

(12) kő tegől zereth. az nem epedh: ha kő keserget akkor wijgad
    who you-Acc. loves that not longs if who moans then rejoices
    ‘those who love you, do not long: if they should moan, they rejoice’

(CzechK. 51–52)

The structure, using the example of hogy kitől in (11), is shown below (see also Bácskai-Atkári 2011: 112–113):

```
CP
 C'
  C_Force
  |  CP
   |  C'
    |  hogy kitől
     |  C_Fin
      |  ...
       |  Ø
```

**Figure 4:** The structure of C_{Force} + operator combinations

Such combinations were not available for mint or mert, which again shows that these could not be C_{Force} heads in complex complementisers, whereas hogy and ha were.

6 Movement and complex complementisers

Besides configurations reflecting the underlying order, there were ones with exactly the opposite surface structure. It has already been mentioned that C_{Fin} heads were ultimately reanalysed as C_{Force} heads. This is possible if the lower C head starts moving up to the upper one, and is finally base-generated there, making movement unnecessary. In clauses where there is only a C_{Fin} head and the C_{Force} head position is empty, it happens in a straightforward way. More interestingly, however, in cases where the C_{Force} head is already occupied by another element, what happens is that the underlying order changes when the C_{Fin} head moves up to be adjoined to the C_{Force} head. This is because of Kayne’s Linear Correspondence Axiom (Kayne 1994), which predicts that an element Y moving up to an element X will be adjoined from the left, resulting in the order YX, which is the mirror image of the underlying XY order.

This process took place in the case of all the combinations discussed in section 5, resulting in the configurations minthogy ‘than that’, mintha ‘as if’, mert hogy ‘because that’ and hogyha ‘that if’. Again, the last one will be discussed in section 7. What is important now is that initially these complex complementisers had exactly the same meaning as the ones reflecting the underlying order.

First, minthogy ‘than that’ was used in comparative subclauses. Consider:

(13) sèmi nagob nem mondathatik: mint hogh legon iftenek
    nothing greater not say-Pass.Cond.3.Sg. than that be-Subj.3.Sg. God-Dat.
    ania
    mother-Poss.3.Sg.
    ‘nothing can be said to be greater than that she be the mother of God’

(TihK. 143)
(14) **mynth hogy ewnnen magat oly allathan enyeztene**

than that PRON.REFL.himself/herself-Acc. such state-Inf. perish-Cond.3.Sg.
‘than (s)he should perish himself/herself in such a state’
(ErsK. 557; ex. from Haader 1995: 563)

The complementiser *mintha* ‘as if’ appeared in conditional comparatives:

(15) *ki meneneć ʒoçafel ʒerent, mint ha aʒ imadʃagə*

out went-3.Pl. custom-Poss.3.Pl.according as if the prayer-Subj.
meneneć
go-Cond.3.Pl.
‘they went out as was their custom, as if going for prayer’
(GuaryK. 113–114)

(16) *aloythwan mynt ha az egheez ʒyghet ʒyngadozna*

thinking as if the whole island-Acc. fluctuate-Cond.3.Sg.
‘thinking as if the whole island had been fluctuating’
(ÉrdyK. 314; ex. from Haader 1995: 543)

The combination *merthogy* ‘because that’ was used in clauses of reason:

(17) **Mert hogjy bizonul ʒuoltualna cristusnak tekelletes**

because that definitely was-3.Sg.be-Cond.3.Sg. Christ-Dat. perfect
tanojītanýa
student-Poss.3.Sg.
‘because he was a perfect student of Christ’
(JókK. 20–21)

(18) **De mer hogjy bodog ferenc zerzetteuala hogy ne**

but because that blessed Francis ordered-3.Sg.was-3.Sg.that not
varnak
wait.-Cond.3.Pl.
‘but because Francis the Blessed had ordered that they should not wait for him’
(JókK. 84; ex. from Haader 1995: 651)

Finally, *hogyha* ‘that if’ was found in conditional clauses – either in ordinary conditional subclauses, as in (19), or as in conditional comparatives, as in (20):

(19) *vig orchaual elmegien vala, hogiha ingen nem*

happy face-Com. away.went-3.Sg.was-3.Sg. that.if absolutely not
hallanaya
hear.-Cond.3.Sg.
“(s)he went away with a happy face, as if (s)he had absolutely not heard it’
(VirgK. 81)

(20) *Es az lattatic ennekom, hoʒ ha az paradičomnak əneʃerufeges*

and that shown-3.Sg. I-Dat. that if the Paradise-Dat.beautiful
edes lakodalmañ lakoznam
sweet dwelling-Ine. live.-Cond.1.Sg.
‘and that is shown me, as if I lived in the beautiful, sweet dwelling of Paradise’
(NagyszK. 118; ex. from Haader 1995: 519)

As can be seen, all of these combinations differ from their counterparts in the previous section with respect to the order of the two complementiser elements: the meaning is always the same.
7 The proposed analysis

Let us now turn to the final analysis for the complex complementisers in question and summarise what has been touched upon so far.

The default word order is basically the one that reflects the underlying structure, hence: \( C_{\text{Force}} + C_{\text{Fin}} \). This naturally gives the configurations of \( \text{hogy} + \text{mint} / \text{mert} \) and of \( \text{ha} + \text{mint} \). The structure is illustrated by the example of \( \text{hogymint} \) below:

![Diagram](image1)

**Figure 5**: The combination of two separate C heads

The reverse order is derived via movement: if the \( C_{\text{Fin}} \) head moves up to the \( C_{\text{Force}} \) head, adjunction will happen in the reverse order, in line with Kayne's Linear Correspondence Axiom (Kayne 1994). This naturally gives the configurations of \( \text{mint} / \text{mert} + \text{hogy} \) and of \( \text{mint} + \text{ha} \). Using the example of \( \text{hogy} \) and \( \text{mint} \), the structure illustrated in Figure 5 is transformed into the one given in Figure 6, resulting in the order \( \text{minthogy} \):

![Diagram](image2)

**Figure 6**: The formation of grammaticalized complex C heads

It has to be mentioned that movement ultimately led to the appearance of grammaticalized complex complementisers, i.e. ones that were already base-generated as a complex \( C_{\text{Force}} \) head. This way there was no need for movement any more, which can easily be explained by the notion of economy: it is more economical to have one grammaticalized complex C head in the grammar than to have movement involved in forming a complex unit.
The analysis presented here has several advantages. First, it is able to capture the reason why the orders XY and YX are just variations of one another, in the sense that the original name is always the same.

Second, it also explains a diachronic difference, namely that the ones with the reverse word order survive into Modern Hungarian, whereas the ones with the default word order have disappeared from the language. This has to do with a change in the constraint on the appearance of two overt C heads. While Old and Middle Hungarian allowed the two C heads to be filled simultaneously, Modern Hungarian does not, as was mentioned before. Hence, the configurations having two separate C heads overtly had to disappear from the language, as opposed to grammaticalized complex C heads, which could remain.

Last but not least, the analysis is also able to explain why configurations (such as mint + mert) are impossible: mint and mert were both CFin heads and it is not allowed to have two of these at the same time, as they cannot both be base-generated in one and the same position.

There remains one problem that has to be accounted for, which is the case of hogy and ha. One might wonder how two CForce heads may appear together if two CFin heads, as has just been established, cannot. The solution lies in the fact that in this case ha was in the upper C head position and hogy in the lower one. Hence, the default word order gives hahogy and so hogyha is the result of movement and inversion. As has already been mentioned, a new operator hogy appeared in the period, which could then easily work in exactly the same way as mint or mert, i.e. become a CFin head and move up. Since hogy was preferred to move to the higher C head, the word order reflecting the underlying structure is relatively infrequent, as compared to hogyha, which survives even into Modern Hungarian.

The fact that hahogy contains two separate C heads is further reinforced by examples where there is a constituent intervening between these two heads:

(21) ha késen hogy el nyugot az nap, hamar esot
if late that PREV set-3.Sg. the sun soon rain-Acc.

‘if the sun has set late, expect rain soon’

(Cis. G3)

As késen ‘late’ can intervene between ha and hogy, it is obvious that they could not form a grammaticalized complex complementiser unit and therefore this is truly an underlying order – which, conversely, disappeared later from the language. In this way, the combinations of hogy and ha fit in the system established for the other complementisers.

To summarise the development of Hungarian complex complementisers, consider Figure 7, the process illustrated by ha + mint:
As can be seen, the development of complex complementisers had altogether four stages. The fourth stage is when the complex unit is fully grammaticalized and only those complementisers that actually reached this level could survive into Modern Hungarian – the others had to disappear from the language.

8 Conclusion

The aim of the present study was to investigate the evolution of Hungarian complex complementisers, focussing on the combinations of the four present-day complementisers with each other. The combinations were found to exhibit a completely symmetrical pattern: if there was a certain combination XY, then the reverse order YX had to exist as well (and vice versa). Moreover, only one member of these pairs did actually survive into Modern Hungarian.

I argued that both of these phenomena are rule-governed and can be explained by the relative cycle, which took place in several constructions in the period, and by Kayne’s Linear Correspondence Axiom, by which an order YX can be derived via head movement from the underlying XY order, thereby explaining the necessary relatedness of such pairs. Finally, I showed that the configurations involving two separate C heads had to disappear from the language later on, as the parametric setting no longer allowed the co-presence of two complementisers in one Left Periphery. By contrast, grammaticalized complex C heads could remain and hence these can be found in Modern Hungarian as well.
References


