

ON THE DIACHRONIC DEVELOPMENT OF A HUNGARIAN DECLARATIVE COMPLEMENTISER*

Julia Bacskai-Atkari

My paper investigates the diachronic development of the Modern Hungarian finite declarative complementiser *hogy* ‘COMP’¹. In Old Hungarian, *hogy* could be combined with other complementisers, e.g. *mint* ‘than/as’, giving configurations like *hogymint* and *minthogy*. That is, complementiser combinations in general are attested both in the *hogy*+X and the reverse X+*hogy* orders, X standing for an unspecified complementiser. The rich variation of Old Hungarian complex complementisers is not fully reflected in Modern Hungarian: it is invariably only one of the orders that survived. I will show that it is always the one that fully grammaticalised into a single C head; this is ultimately tied to the original underlying order of *hogy* and X as separate C heads. I will also demonstrate that *hogy* came to be used as a general marker of finite subordination.

1. THE THEORETICAL BACKGROUND

The questions outlined above will be approached from a minimalist perspective. In current minimalist approaches, the structure of a clause can be divided into a thematic domain, the VP (verb phrase) and a functional domain, consisting of the TP (tense phrase) and the CP (complementiser phrase). All of these are referred to as domains because all of them may contain several positions of the same type, hence there can be multiple verbs or complementisers in a single clause.

* 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). I owe many thanks to Philomen Probert and Andreas Willi for their helpful remarks on the earliest version of this paper.

¹ The closest equivalent to *hogy* in English would be *that*; however, I will gloss *hogy* as ‘COMP’ because the environments in which *hogy* is (and was) licensed to appear are slightly different from those in which the Modern English *that* does. Rather, I would like to indicate that *hogy* is a general subordinator, and the specific syntactic environments will be specified when necessary.

The structure of phrases can be represented using the X'-schema given in *Figure 1*²:

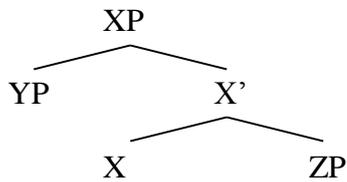


Figure 1

The CP domain, also referred to as the left periphery (or left edge), is also built up of such XPs, and it is present both in main and in embedded clauses: it is responsible for defining the Force of the clause, i.e. whether it is declarative, interrogative etc. In embedded clauses, it contains subordinating conjunctions and various *wh*-elements. Following Rizzi (1997), the structure of the left periphery contains two CPs,³ as shown in *Figure 2*:

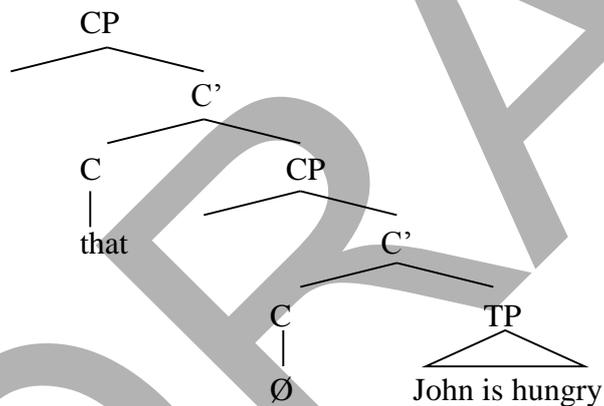


Figure 2

The higher C takes another CP in its complement position; the head of that lower CP takes the rest of the clause (given here as TP, e.g. *John is hungry*) as its complement. Embedded clauses can be introduced complementisers, which roughly correspond to subordinating

² Disregarding now the exact mechanism of how syntactic structures are built up, the schema given in *Figure 1* can be applied to any phrase XP. Every phrase has a head (X), which can take a specifier (YP) and a complement (ZP). While the specifier and the complement are full phrases, the head is not.

³ Note that Rizzi (1997) uses a more complex CP-system, in which other projections (TopP, FocP, etc.) may appear in between the two CPs hosting complementisers. This is not relevant for the present discussion, since I am largely concerned with the combinations of complementisers only, and Hungarian (both now and historically) rarely exhibits high topics, as topics are normally found below the CP-domain (and focus always is).

conjunctions, such as *that*, *if* or *than*. These may take either the higher or the lower C head position: for instance, English *that* is a higher C head, see *Figure 2*. While in some languages always only one of the C heads is filled by a complementiser (e.g. Italian, cf. Rizzi 1997), in others it is possible that two overt complementisers co-occur (e.g. Welsh, cf. Roberts 2005: 122).⁴

Complementisers are base-generated in a C head position: this means that when the clause is constructed in a bottom-up fashion, they are inserted into this position (depending on the function of the given complementiser, this may either be a higher (Force) or a lower (Fin) C head). The operation responsible for combining any two syntactic elements is referred to as Merge: taking the example in *Figure 2*, Merge combines the TP (*John is hungry*) with the lower C head – ultimately to form the (lower) CP projection. Next, Merge combines this lower CP with the higher C head *that* to form the higher CP projection.

Besides complementisers, (relative) operators can introduce subordinate clauses, in the sense that they appear as the first elements in the subordinate clause. Consider the following examples (the symbol % in (1e) indicates that the acceptability of the sentence shows dialectal/idiolectal variation):

- (1)
- a. I don't remember **who** wrote the book.
 - b. I don't remember **what** he wrote.
 - c. I don't remember **which book** he wrote.
 - d. I don't remember **when** he left.
 - e. % John is taller than **what** Mary is. (Chomsky 1977: 87, ex. 51a)

⁴ The reason why there are two distinct C heads is, according to Rizzi (1997), that they have different functions. The lower C head is responsible for defining the finiteness of the clause: while finite clauses contain a tensed verb (e.g. *John is hungry* is a finite clause), non-finite ones do not (e.g. the clause *to go to Berlin* is a non-finite one in a complex sentence such as *I want to go to Berlin*). Higher C heads, on the other hand, are responsible for defining the Force of the clause, i.e. whether it is declarative, interrogative, relative etc.

The elements given in boldface in (1) are all relative operators. They are syntactically different from complementisers in several respects. First, they are phrase-sized constituents, as is clearly indicated by (1c), where the determiner *which* takes a lexical (nominal) complement (*book*). Second, they also function as arguments or adjuncts in the clause that they introduce: *who* in (1a) is a subject, *what* and *which book* in (1b) and (1c) are objects, *when* in (1d) is a time adverb (hence an adjunct), while *what* in (1e) is a nominal predicate. Since operators are phrase-sized, they cannot be C heads: instead, they occupy the specifier position of a CP (the head of which is zero in these cases). In addition, they are not base-generated in the CP-domain (unlike complementisers) but they are moved to this position via *wh*-movement (cf. Chomsky 1977: 87; Kennedy & Merchant 2000: 89–90).⁵

It has to be mentioned that languages may have different settings as to which CP will ultimately host the moved element. Let us take the example in (1e), which is a comparative subclause: *what* follows the complementiser *than*, which is located in the higher C head position, hence *what* must be in the specifier of the lower CP. The representation is shown in

Figure 3:

⁵ The notion of movement in generative grammar serves to relate two main functions that *wh*-elements fulfil: they establish semantic relationships within the clause just as their non-*wh* counterparts do but at the same time they also serve as elements introducing the subclause. Taking the example in (1c), the *wh*-element *which book* is clearly the object of the verb (*wrote*). Objects normally follow the verb (e.g. *he wrote a book*): their base-position is within the lexical VP. Since *which book* in (1c) is an object, it is expected to be base-generated in the same position as *the book*. However, in (1c) it appears in a different position in the final structure, which is the same position where we find *wh*-elements in main clause questions too (e.g. *which book did he write?*). Movement is a way to overcome the apparent contradiction of having an element in two positions: *which book* originates (i.e. is base-generated) in the VP but then it moves to the CP-domain, i.e. to the specifier position of the CP.

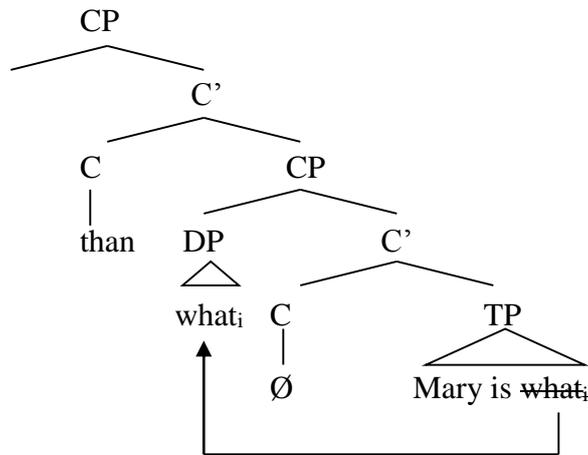


Figure 3

As can be seen, the *wh*-element is a DP (determiner phrase, a functional extension of the nominal expression) that originates in the TP but moves up to the specifier of the lower CP.

There are thus two identical copies of *what*: however, only the higher one is pronounced. The DP *what* is inserted into the structure for the second time via Merge: hence in the case of movement, an element is first merged into the structure in its base position, and subsequently moves up to a different position to be merged into the structure again. Note that it is far simpler to have complementisers in the CP-domain as they have to be merged only once, cf. Chomsky (1995). However, the option of direct Merge would not be available for operators that have other functions to fulfil within the TP domain (see above).

Of course, there are several other theoretical issues that could be addressed; however, what is important for the present discussion is that though both complementisers and operators may introduce subordinate clauses,⁶ they take syntactically distinct positions and hence exhibit different syntactic behaviour – moreover, they may also co-occur in a predictable way.

⁶ Note that, as was discussed earlier in detail, by saying that an element can introduce a subclause I do not intend to say that the particular element would be the head of the clause as this is clearly not the case for operators, which are distinct from C heads. What is common in complementisers and relative operators is that both can appear clause-initially and as such may serve as the overt markers of a clause being subordinate – however, syntactically it is naturally only a C head that may actually determine the subordinate nature (and the specific Force or Finiteness) of the clause.

Modern Hungarian also allows overt higher C heads, operators, and occasionally the co-occurrence of the two. Consider the following examples:

- (2) a. Tudom, **hogy** Péter alszik.
know-1.Sg. COMP Peter sleeps
 ‘I know that Peter is sleeping.’
- b. Akkor indulok, **amikor** Péter.
then start-1.Sg. when Peter
 ‘I will leave when Peter does.’
- c. Gyorsabb vagyok, **mint amilyen** Péter.
faster am than how Peter
 ‘I am faster than Peter.’

In (2a), there is a single C head *hogy* ‘COMP’, which occupies the higher C head position in the subordinate clause (cf. É. Kiss 2002: 230–242; Kenesei 1992). In (2b), there is a single relative operator, which is located in the lower [Spec,CP] position (cf. Kántor 2008). Finally, in (2c) there is a combination of the higher C head *mint* ‘than’ and a comparative relative operator, *amilyen* ‘how’, which is in the lower [Spec,CP] position (Kántor 2008; Bacskai-Atkari 2012; on the status of *mint*, cf. also Kenesei 1992). These positions correspond to the ones indicated in *Figure 2* and *Figure 3* for English and hence the CP-domains of the two languages are essentially similar in this respect. Note that the structure of the Hungarian clause is noticeably different as far as various further left-peripheral projections are concerned (cf. É. Kiss 2002 for a detailed analysis) but these projections are irrelevant for the present discussion.

One of the most intriguing questions concerning the diachronic development of complementisers and operators is precisely whether and to what extent they are related: in other words, whether a given element that belongs to one category now used to belong to another and if so, how the change between the two positions can be conditioned. Section 4, will discuss how complementisers develop from operators; before turning to this, let us first consider the Hungarian data to be accounted for.

2. CHANGES IN HUNGARIAN DECLARATIVES – AN OVERVIEW

In Modern Hungarian, the complementiser *hogy* ‘COMP’ is responsible for introducing finite clauses, similarly to *that* in English. In Old and Middle Hungarian, *hogy* had a wide range of functions: for instance, it appeared in comparative clauses as well, while Modern Hungarian has a complementiser for this functions (*mint* ‘than/as’).

Furthermore, the issue of multiple and complex complementisers must also be addressed. The notion of multiple complementisers mean that there are two separate complementisers in a clause (hence both C heads are filled, see section 1); complex complementisers refer to configurations with a single complementiser head that is morphologically complex, i.e. it is a conglomeration of two once separate complementisers. Historically, there are several multiple and complex complementisers in Hungarian: many of them have not survived into Modern Hungarian, such as *hogymint* ‘COMP than’, while others still exist, such as *minthogy* ‘than COMP’.⁷

There are two main points of interest here: first, if a combination existed in a *hogy*+X order, (X standing for another complementiser), then there was also an X+*hogy* reverse order combination (originally with the same meaning), and the existence of an X+*hogy* combination implies the existence of a *hogy*+X combination. Second, it is only one of the orders that survived: the other one invariably disappeared from the language before the Modern Hungarian period.

The possible combinations of complementisers are summarised in *Figure 4* (the combinations *hamint* ‘if as’ and *mintha* ‘as if’ are in parenthesis because they do not contain *hogy* that is

⁷ Note that, for the sake of clarity, in translating complementiser combinations, I use a morpheme-by-morpheme translation as providing merely the meaning of the whole combination would result in a loss of grammatical information, e.g. both *mint* and *hogymint* meant ‘than/as’ but for the argumentation to be pursued here, it is vital to see the morphological structure of such complexes.

under scrutiny here; nevertheless, they conform to the general system established for the ones containing *hogy*):

	ha 'if'	hogy 'COMP'	mert 'because'	mint 'than/as'
ha 'if'	–	hahogy	–	(hamint)
hogy 'COMP'	hogyha	–	hogy mert	hogymint
mert 'because'	–	merthogy	–	–
mint 'than/as'	(mintha)	minthogy	–	–

Figure 4

As can be seen, for any of the complementiser combinations given in the chart, it is true that it existed in both possible orders: however, it is only one of them that survived (highlighted in Figure 4).⁸

Apart from combinations with other complementisers, *hogy* also appeared in relative clauses: co-occurrences with ordinary relative operators such as *ki* 'who' and *mi* 'what' are attested in Old and in Middle Hungarian, resulting in sequences such as *hogyki* 'COMP who' and *hogymi* 'COMP what'.

In the following, I will try to account for how the functional changes concerning *hogy* can be explained and how its co-occurrences with other elements in the CP-domain are conditioned.

All the changes and differences will be linked to the changes in the structural positions of

⁸ Note that not all logically possible combinations exist or existed and this largely has to do with the semantic properties of the given complementisers: certain combinations are semantically incompatible, such as the combination of *mint* 'than/as' with *mert* 'because'. As will be shown later on, it is precisely the semantically vacuous nature of *hogy* 'COMP' which allowed it to appear in a large number of combinations.

hogy and of the other complementisers. I will show that *hogy* – just like all the other complementisers – developed via the relative cycle, as described by van Gelderen (2009); in addition, I will demonstrate that it became a general marker of declarative Force in Old and Middle Hungarian – hence its strong potential for combining with other elements.

My account is strongly based on Rizzi's (1997) model of the left periphery on the one hand, and on generally attested grammaticalisation processes on the other hand; these enable one to understand the systematic diachronic changes behind the synchronic and diachronic Hungarian facts. The advantage of the proposal is that it may account for these phenomena as parts of a system, instead of providing partial analyses; furthermore, the application of cross-linguistically attested mechanisms also relate the Hungarian changes to more general processes. In this way, the present account is strongly restrictive in terms of what may qualify as a possible grammaticalisation process, which also increases the explanatory force of the analysis.

3. METHODOLOGY AND AIMS

The existence of the various combinations of *hogy* 'COMP' is well-known in the traditional descriptive historical grammars on Hungarian. The importance of the present study lies in providing a formal account of these combinations and to show that complementiser combinations did not appear at random but there are important regularities to be observed (for instance, the symmetric nature of the combinations, as given in *Figure 4* and the predictability of which combinations survive based on general principles). This also means that there are structural differences among the various combinations that are not noted in traditional

grammars.⁹ Last but not least, the role of *hogy* in these combinations also needs to be clarified.

Apart from building on existing descriptive findings, my study also includes corpus search. To date, there is only one searchable historical corpus on Old Hungarian, started by the “Hungarian Generative Diachronic Syntax” Project (2009–2013, OTKA-78074): this is the “Old Hungarian Concordance” (2.2 million tokens; English version available under: <http://omagyarkorpusz.nytud.hu/en-search.html>; metadata such as tokens and dates are also taken from here), which includes all extant records from the Old Hungarian period (for more information on the corpus and the project, cf. Egedi & Simon 2012). Since, to date, contains most of the texts only in the original spelling, quantitative research is severely limited, as there is considerable variation in Old Hungarian spelling: for instance, there are at least 8 spelling variants for *hogy* itself (disregarding initial capital letters and punctuation marks directly following the last letter). The number of variants increases for combinations and since only a very small part of the corpus is morphologically annotated, the results of a single search must be checked by hand.

In addition, a proper comparison with Middle Hungarian data is not available either, there being no available searchable corpus of Middle Hungarian texts at all. Considering all these factors, for the time being I cannot present reliable quantitative data of all the occurrences of the combinations. On the other hand, the fact that new examples (that is, examples that have not been discussed in the existing literature) could easily be found even in earlier texts

⁹ Most of the literature concerning the syntax of earlier stages of Hungarian is purely descriptive and does not intend to provide a formal account for the diachronic changes, let alone to investigate the specific syntactic positions associated with the various elements. The culmination point of this kind of literature is the three-volume work entitled *A magyar nyelv történeti nyelvtana* [The diachronic grammar of the Hungarian language], issued in 1991, 1992 and 1995 – in the present article, I refer to the chapters Haader (1991, 1995) and Juhász (1991, 1992).

suggests that complementiser combinations were well-established in Old Hungarian as well, and this refutes the idea such combinations were insignificant in Late Old Hungarian.

4. THE RELATIVE CYCLE

The relative cycle is a grammaticalisation process, whereby an original demonstrative or interrogative pronoun becomes first an operator moving to [Spec,CP], and subsequently this operator is reanalysed as the head of that CP; in turn, the operator position becomes available for new operators (van Gelderen 2004, 2009; Roberts & Roussou 2003). Since in a split CP relative operators are located in the lower CP (see section 1), the new complementiser is initially a lower C head; later on, it can be reanalysed into a higher C, which is the case for English *that*, as shown by van Gelderen (2009) and also for German *dass* ‘that’, as shown by Axel (2009). The lower C position hence represents an intermediate stage between pure operator and pure complementiser status, and can host C heads that have not yet been completely grammaticalised (see Bayer & Brandner 2008 on the intermediate status of word-sized C elements); the higher C head is responsible for marking subordination and Force, and hence properly grammaticalised complementisers are expected to appear in the higher C in a split CP. The major change constituting the reanalysis from operator into C head is motivated by economy: economy is a basic principle of generative grammar, which ensures that the derivation of syntactic structures (the way clauses and phrases are built up in a bottom-up fashion) is as simple as possible and includes only a minimal number of steps. Recalling what was said in section 1, it should be clear that single merge is preferable over movement (which involves the merge and re-merge of the same element). This is formalized in two main principles: the Head Preference Principle (HPP) and the Late Merge Principle (LMP), cf. van Gelderen (2004).

The HPP states that it is preferable to be a head than a phase; the LMP states that it is more economical to be base-generated in a higher position than to be moved to that position. Both indicate that it is more economical to be base-generated as a C head in the CP-domain than to be moved there as an operator. Naturally, this reanalysis is possible only if the elements in question lose their functions that they used to have in the TP-domain: if this condition is met, elements are likely to be reanalysed as functional C heads.

I claim that Hungarian *hogy* ‘COMP’ underwent the relative cycle, and was accordingly reanalysed from a relative operator into a C head; in addition, its position as a C head also changed from lower into higher C. The processes summarised in *Figure 5* refer to the changes of Hungarian *hogy*:

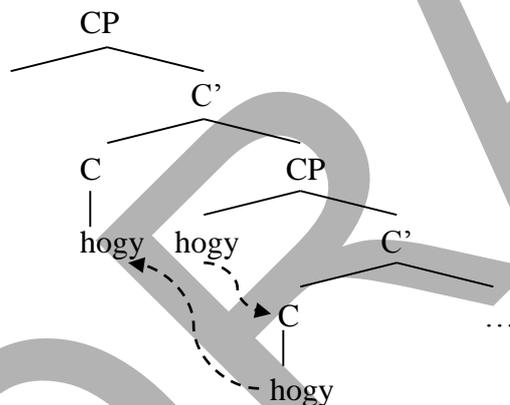


Figure 5

As can be seen, *hogy* first appears in the specifier position of the lower CP, which is the designated position of operators (cf. section 1). As a next step, *hogy* is reinterpreted as the head of the same (lower) CP. Finally, with its establishment as a complementiser, it is reanalysed from a lower to a higher C head. The various functions of *hogy* can be found in the earliest texts, indicating that the processes had begun before Old Hungarian, and hence Old Hungarian saw the completion of the cycle. These functions are illustrated in (3) below: in

(3a), *hogy* is an interrogative operator, in (3b) it is a relative operator, and in (3c) it is a complementiser:

- (3) a. **Hogý** kynozá ewnmagat poclosual egý talbalol euen
how tortured-3Sg himself-Acc sinner-Instr one plate-Ela eating
 ‘how did he torture himself by eating with a sinner from one plate?’ (*JókK.*, 101)
- b. furifcte mufia!|| etetý ýmletí. ug **hug** ana fciuluttet.
bathes washes feeds breastfeeds so how mother child-Poss.Acc
 ‘she bathes, washes, feeds and breastfeeds him as a mother does her child’
 (*KTSz.*)
- c. wylagnoc kezdetuitul fugua? rohtonc ez nem levt
world-Dat beginnings-Poss.Abl from we.Sup this not became.3Sg
 wala. **hug** fcu3 lean. fiot fciulheffen
be.Aux COMP virgin girl son-Acc bear-Possib.3Sg
 ‘we have not experienced this from the beginning of the world, namely that a son could be born to a virgin girl’ (*KTSz.*)

The interrogative operator and the complementiser use of *hogy* survive into Modern Hungarian. Regarding the relative operator, the form *ahogy* ‘how-Rel.’ can be detected: while relative operators in early Old Hungarian had an identical morphophonological shape to their interrogative counterparts, a distinctive form with *a-* started to appear in late Old Hungarian (see section 4; cf. Sipos 1991: 398, G. Varga 1992: 524–525), which eventually became obligatory for all relative operators. Hence the use of *hogy* in (3b) is on the one hand continued by *ahogy*, and on the other hand it led to the appearance of constructions like (3c)

via the relative cycle, and the resulting complementiser was naturally unaffected by the morphophonological changes applying to relative operators.⁵ **SIMPLEX COMPLEMENTISERS**

The other present-day Hungarian complementisers also developed by way of the relative cycle; these are: *ha* ‘if’, *mint* ‘than/as’ and *mert* ‘because’. All of these, as well as the declarative *hogy* ‘COMP’, were originally pronouns, which came to be operators (cf. Juhász 1991: 479–481, 1992: 781, 783–785, 801; Haader 1991: 729–737, 1995: 510–677): *hogy* meant ‘how’, *ha* meant ‘when’, *mint* meant ‘how’, and *mert* meant ‘why’.¹⁰ Due to a functional split between the original operator and the newer complementiser functions, the related (interrogative/relative) operators can still be found in the language.¹¹ Ultimately, the changes led from a configuration in (3), whereby all the three functions are essentially surface-identical (disregarding spelling variations), into a configuration where the individual forms are specialised for functions. In the case of *hogy*, this means the following: complementiser is established as *hogy* (no variation), the interrogative operator is *hogy/hogyan* ‘how’ (but note that interrogative operators bear focal stress in Hungarian, unlike C heads or relative operators, hence there is prosodic distinction also for the form *hogy*), and the relative operator is *ahogy/ahogyan* ‘how’ (hence clearly marked as a relative operator due to its initial *a*-vowel).

¹⁰ For an early example involving the operator use of *hogy*, see example (3b) in the previous section.

¹¹ Operators generally develop into C heads introducing clause types in which they are licensed as operators. This does not exclude the possibility of later functional changes; however, there are certain typical co-occurrences, e.g. a temporal operator (e.g. *when*) is frequently reanalysed as a conditional, or a modal operator (e.g. *how*) as a comparative complementiser. While operators have to lose features incompatible with C heads when they grammaticalise into C heads, this does not apply to features that are fundamentally related to clause-typing. Hence most ordinary relative operators cannot be reanalysed into C heads in Hungarian because Hungarian does not permit the presence of, for instance, person and number features (or overt case marking) on C heads; however, a comparative operator equipped with the features [+rel] and [+compr] (relative and comparative) could be reanalysed into a C head also marked as [+rel] and [+compr]. For a general subordinator like *hogy*, the loss or absence of very specific clause-typing features is also necessary, and *hogy* was an ideal candidate for two main reasons. First, as an operator it was usually an adjunct, hence an element only loosely associated with the argument structure of the subclause, and as a VP-adverb it did not bear case-endings or person and number markers either. Second, the subclauses containing *hogy* were not necessarily taken by a specific element in the matrix clause that would unambiguously require a specific kind of subclause (as, for instance, the [+compr] degree element *-er* in a comparative matrix clause selects for a [+compr] subclause); hence these environments also favoured the reinterpretation of *hogy* as a general subordinator. These conditions

The syntactic differences between complementisers and operators were outlined in section 1 in detail: a complementiser is a head while an operator is phrase-sized; consequently, lexical material may appear together with an operator but not with a complementiser. As the reanalysis of an operator into a C head is contingent upon the absence of lexical material, it is expected that during the period of change there are examples that are ambiguous between a C head and an operator. Hungarian shows a strong prohibition on inflectional morphemes appearing on operators, unlike languages that show complementiser agreement, cf. Bayer (1984), Bayer and Brandner (2008) on Bavarian, and Willis (2007) on Welsh. Moreover, if another operator is allowed to co-occur with a given element in one left periphery, the latter cannot be an operator any longer since then the two would be competing for the same position – this is shown for the reanalysis of *mint* and the appearance of new comparative operators in Hungarian in Bacskai-Atkari (2011, 2014) in detail. Furthermore, there is also an important morphological requirement on Hungarian relative operators that started in Middle Hungarian:¹² interrogative operators kept their original forms (e.g. *ki* ‘who’), while relative operators developed a specific form starting with *a-*, hence an original relative operator *ki* ‘who’ developed into *aki* ‘who’.¹³ This is not shown by the four complementisers, unlike their

were not met by other operators in Hungarian. On the other hand, it is worth mentioning that the reinterpretation of an operator *how* into a declarative complementiser is attested in other languages as well: for instance, this change has partially taken part in the case of *how* in contemporary English, as described by Willis (2007: 434) – see also van Gelderen (2009: 144–145).

¹² The three main periods of the Hungarian language are traditionally as follows: Old Hungarian (896–1526), Middle Hungarian (1526–1772) and Modern Hungarian (from 1772 onwards).

¹³ The origin of the prefix-like *a-* is a demonstrative pronoun in the matrix clause (*az* ‘that’). The role of this matrix pronominal element is to mark the syntactic status of the subclause with respect to the matrix clause: it takes on the relevant case endings, and it appears in a position that shows the information structural status of the subclause (e.g. topic, focus), see É. Kiss (2002: 244). The pronoun may stand by itself and it may also co-occur with a lexical head. Configurations in which there was no lexical (or any other) element intervening between the matrix pronominal element and the relative pronoun lead to the reinterpretation of the status of the matrix pronoun: it cliticised onto the following element anyway, and it started to be reinterpreted as part of the relative pronoun, accompanied by a gradual phonological reduction, as described recently by Bacskai-Atkari (2014a: 270–275), following É. Kiss (2002: 243–244) and Haader (2003: 507). Hence the morphophonological change from an initial string *az ki* ‘that who’ can schematically be illustrated as follows: *az ki* > *azki* > *akki* > *aki*. Note that these processes started in late Old Hungarian already and continued into Middle Hungarian: in other words,

relative operator counterparts, which are *ahogy* ‘how’, *amint* ‘how’ and *amiért* ‘why’ (note that the operator *ha* ‘when’ was rare already in Old Hungarian and disappeared during Middle Hungarian altogether).

Regarding the feature changes for *hogy*, the following can be observed. As an interrogative operator, *hogy* was naturally equipped with a [+wh] feature; as a declarative complementiser, it is expected to have lost this feature specification. This is indeed the case but note that *hogy* is not a [-wh] element either: rather, it is insensitive to [\pm wh] and may hence introduce both [+wh] and [-wh] embedded clauses (see É. Kiss 2002: 239). This can be detected already in Old Hungarian for the complementiser *hogy*: the example in (3c) in section 4 also demonstrates the use of *hogy* in a [-wh] embedded clause. The examples in (4) show that *hogy*-clauses can be complements to verbs that select for [+wh] clauses, such as *kérdez* ‘ask’ in (4a), and also to verbs that select for [-wh] clauses, such as *mond* ‘say’ in (4b):

- (4) a. Azert immar kerdi zent Bernard doctor ezt **ho**g mel’ic
therefore now asks saint Bernard doctor this-Acc COMP which
 tezon na’gob bint ek ketto kozzul
makes greater sin-Acc this two among
 ‘therefore asks Saint Bernard doctor, which of the two constitutes a greater sin’
 (*GuaryK.*, 25)
- b. De v’g mond zent Bernald, **ho**g v’g kell embernek tartanja
but so says saint Bernard COMP so must human-Dat hold-Inf.3Sg
 az o’ testet, mikeppen az oruos tart’a a kort
the his body-Acc how the doctor holds the illness-Acc
 ‘but Saint Bernard says that people must tend to their bodies as a doctor tends to
 sickness’ (*BodK.*, 5v)

Though the processes are very much alike for all the four complementisers, there is an important difference with respect to the chronology. In the case of *hogy* and *ha*, the functional split took place before the Old Hungarian period. In Old Hungarian *ha* was always in the

the appearance of the modern relative operators in Middle Hungarian simply means that they were established as grammaticalised entities in that period, but the grammaticalisation processes producing them naturally started earlier.

higher C head position, while *hogy* was typically a higher C head and rarely a lower one: this has to be so if elements start to be responsible for marking the Force of the clause, which is associated with the higher C head position.

Evidence for the relative (higher or lower C) position of a complementiser comes from the way it combines with other CP-peripheral elements, including other complementisers, polarity heads, and relative operators. As relative operators are located in the lower [Spec,CP] position, only higher C heads may precede them: this is attested for *hogy* (see section 6). The example in (5a) shows *hogy* in comparatives expressing inequality: the element *mint* ‘how’ (either still an operator or a complementiser) is located in the lower C, while the element *nem* ‘not’ is a Pol (polarity) head that regularly occurs between the two CPs in Old Hungarian comparatives (see Bacskai-Atkari 2014b: 489–493; cf. also Bacskai-Atkari 2011 for *hogy* in comparatives). Since *hogy* in (5a) precedes both the Pol and the lower C head, it can only be a higher C head. Examples for a lower C *hogy* in a split CP are scarce: one such example is (5b), where the complementiser *ha* and also a high topic precedes *hogy*, indicating that it can only be in the lower C (it is also clearly not an operator meaning ‘how’); the combinations with *ha* will be discussed in section 8.¹⁴ Consider:

- (5) a. Ionkab adna embor mağat mindon kenokra:
rather give-Cond.3Sg human oneself-Acc all sufferings-Sub
hodnem mint: Cak eđ zompillantasik latna az v
COMP-not how only one moment-Term see-Cond.3Sg the their
zynoket
faces-Poss.Acc
 ‘one would rather give oneself for all sufferings than to see their faces for a moment’ (*KazK.*, 3r)

¹⁴ Note that C + a topic + *ha* combinations did not occur, which indicates that *ha* was no longer base-generated as a lower (Fin) C head. Since the aim of the present article is not an investigation of the diachronic changes in the status of *ha*, I will not elaborate on these issues in more detail.

- b. Ha késen **hogy** el nyugot az nap, hamar esot váry
if late COMP off set-3.Sg the sun soon rain-Acc expect-Imp.2.Sg
 ‘if the sun has set late, expect rain soon’ (Cis., G3) By contrast, for *mint* and

mert the split took place only during Old and Middle Hungarian: as a consequence, *mint* and *mert* were initially located in the lower CP.¹⁵ Hence, for instance, the forms *mert* and *miért* could be used interchangeably both for the operator ‘why’ and for the complementiser ‘because’, while in Modern Hungarian *miért* is always ‘why’ and *mert* is ‘because’. Likewise, *mint* was in free variation with operator elements such as *miképpen* ‘how’ or *miként* ‘how’ only as long as it could function as an operator but this is not permitted e.g. in Modern Hungarian, where it is clearly a C head.¹⁶ Evidence for whether a given complementiser occupied the higher or the lower C head comes from its relative position with respect to other complementisers in combinations, as will be demonstrated in the next section.

The possible Old Hungarian positions for present-day complementisers are summarised in

Figure 6:

¹⁵ This is indicated for instance by the fact that *mint* and *mert* were not allowed to be followed by relative operators, unlike *hogy* and *ha*.

¹⁶ Since the focus of the present study is not the development of either *mint* or *mert*, I do not wish to provide a detailed description of how the changes from operator to head can be observed in these cases or to try to associate the given stages with narrow time periods as either of these would require an investigation far beyond the scope of my investigation here and since neither of these issues is directly linked to the development of *hogy*, they are not of crucial importance here. As far as *mint* is concerned, cf. Bacskai-Atkari (2011).

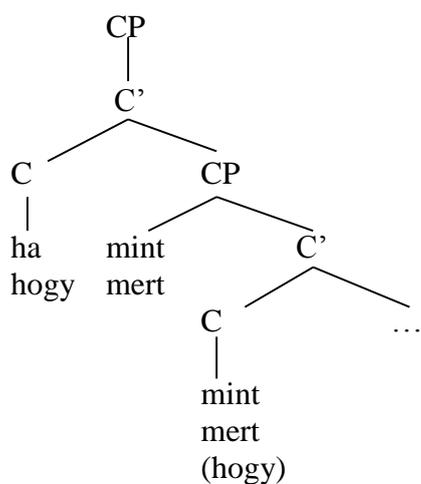


Figure 6

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.¹⁷

6. MULTIPLE COMPLEMENTISERS

Given the mechanism of the relative cycle (see section 4), it is expected that the grammaticalisation of *hogy* into a complementiser makes the original operator position available for new operators. This is indeed the case: *hogy* in Old Hungarian could be followed by relative operators and also by operators that were reanalysed into lower C heads (*mint*, *mert*). Such combinations are illustrated in (6):¹⁸

- (6) a. 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., 141)

¹⁷ Apart from losing their ability to follow higher C heads, this is also indicated by the fact that e.g. *mint* is able to combine with operators in the lower [Spec,CP] position in the order *mint* + operator, which would clearly not be possible if it were a lower C head.

¹⁸ As of 3 April 2015, the normalised part of the Old Hungarian Concordance corpus contains 4 instances of the sequence *hogymint* and 2 instances of the sequence *hogymert*, in altogether 4 codices (60560 tokens altogether; the token number refers to the normalised version without punctuation marks). The example in (6c) is taken from this search; the remaining examples in (6) were searched manually based on the original spelling.

- b. *mínd anne bosegos koñhullatasoc mene a vízeknec*
all so.much plentiful crying-Pl. as.much the waters-Dat.
sokassaghí sem volnanac en elottem kellemetosek/
multitude-Poss.3.Pl. neither be-Cond.3.Pl. I before.me pleasant-Pl.
*Auag foganatosoc **hog** **mínt** akkí zonetlen a kereztfanac o*
or effective-Pl. COMP as who incessantly the rood-Dat. he
keserúseget v̄ testeben v̄iselí
bitterness-Poss.3.Sg.Acc. he body-Poss.3.Sg.Ine. bears
 ‘not even as much crying as the multitude of waters would be as pleasant and touching to me as the one who incessantly bears the bitterness of the rood in his body’ (*NagyszK.*, 40–41)
- c. *Dehogÿ mert zent ferenc ygen zeretuala ewtett*
but.COMP because saint Francis well liked.was-3.Sg. him-Acc.
tÿztasagert es alazatosagaert kyt valuala
purity-Fin. and humility-Poss.3.Sg.Fin. who-Acc. have-3.Sg.was
Monda nekÿ
said-3.Sg. him-Dat.
 ‘but because Saint Francis liked him well for his purity and for his humility that he had, he said to him’ (*JókK.*, 46)
- d. *De **hogy** **meyerth** dichewlth tefbe wagyok en Nen*
but COMP because redeemed body-III. am I not
fyrhatok
can.weep-1.Sg.
 ‘but because I am in a redeemed body, I cannot weep’ (*AporK.*, 158)

The examples above contain subclauses introduced by *hogy* ‘COMP’ and *mint* ‘than/as’ or *mert* ‘because’. Note that I retained the original spelling and hence the fact that two words are written without a space is only because in Old Hungarian orthography there were no set rules as to what was written together and what was not. Hence the lack of a space between two elements does not indicate that they are to be understood as one complex unit, since this was not the case.

The possible structures of the relevant left peripheries in (6) are shown in *Figure 7*:

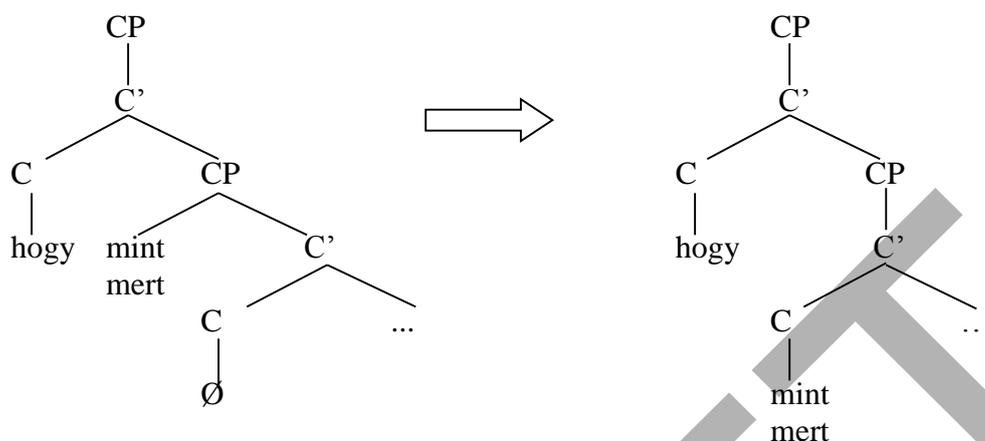


Figure 7

The diagram on the left shows the earlier stage when *mint* and *mert* were operators moving to the specifier of the lower CP: after they grammaticalised into C heads, the relevant structure is the one on the right. Naturally, such configurations had fixed word order predictably conforming to Figure 6: since typically *hogy* was located in the higher C head position, most combinations of two separate C heads (or of a higher C head and an operator) are in the *hogy+X* order – hence the configurations *hogymint* ‘COMP than’ and *hogymert* ‘COMP because’. Note that besides the order of the two elements being constrained, so is the meaning: for all combinations *hogy+X* the meaning is invariably that of X.

As was stated above, all the four complementisers have been reanalysed ultimately as higher C heads by the Modern Hungarian period: none of them is base-generated in the lower C anymore, and hence the C+C combinations should no longer be possible in Modern Hungarian – in other words, if they existed, they could not be combinations involving two separate C heads. Since the C+C combinations *hogymint* and *hogymert* are indeed no longer attested, there is no discrepancy between the analysis and the historical data. In what follows I am going to present further reasons why the particular analysis should be maintained and how it works in detail.

Before doing so, let me mention a logically possible option that one may find tempting to consider as an alternative analysis of sequences such as *hogymint*, but which is clearly untenable both on theoretical and empirical grounds. What I have been arguing for is that *hogy* is a higher (Force) C head and *mint* is either in the lower [Spec,CP] or a lower (Fin) C head and as long as one maintains that *hogy* is a complementiser and that *hogy* and *mint* are syntactically distinct, this is the only possible analysis. The question is whether it is possible that *hogy* was in a [Spec,CP] position and *mint* was the C head of the same projection, that is, as an instance of a Doubly Filled Comp.

This is, however, clearly untenable for a number of reasons. First and foremost, such a configuration would by definition rule out the appearance of any element between *hogy* and *mint*; however, in Old and Middle Hungarian it was frequent for a negative element – such as *nem* ‘not’ – to appear between *hogy* and *mint*, resulting in combinations such as *hogynemmint* ‘COMP not than’ (see (5a); cf. Bacskai-Atkari 2011, 2014). Second, the violation of the Doubly Filled Comp Filter is not attested in Hungarian subordinate clauses elsewhere even in earlier periods and treating *hogy* as an operator in such constructions would raise the problem of how structurally parallel relative clauses (that is, combinations such as *hogy ki* ‘COMP who’) should be analysed since *hogy* is clearly not an operator in these cases. In other words, the claim that *hogy* should be treated as an operator in combinations like *hogymint* would introduce an exception unattested in other constructions. Last but not least, while *mint* unquestionably became a C head, it could initially alternate with *miképpen* ‘how’ and *miként* ‘how’ (cf. Haader 2003: 539), which are clearly operators; these could likewise appear in sequences such as *hogy miképpen* and *hogy miként* but there is no reason to believe that they would be C heads in any period.

7. COMPLEX COMPLEMENTISERS

Let us now turn to the problem of movement. As was argued for in sections 5–6, all present-day Hungarian complementisers went through the relative cycle and ended up in the higher C position, associated with Force. This means that also *mint* ‘than/as’ and *mert* ‘because’, which are in the lower C head in *Figure 7*, started to be reinterpreted as base-generated complementisers and were no longer associated with their original operator functions (completion of the relative cycle).¹⁹ Accordingly, in a split CP they were to be located in the higher C head position. If the higher C head was generated above the one containing *mint/mert* to mark subordination separately, the lower C heads moved up to their preferred peripheral position and joined *hogy* via head movement. The leftmost (highest) position is associated with marking Force, which is carried by *mint/mert* in these configurations, while *hogy* increasingly marked subordination only.

In what follows I will argue that this was how grammaticalised complex complementisers developed: they stem from the sequence of two separate simplex complementisers by way of the lower C head moving up to the higher one and adjoining to it. Adjunction for heads roughly means that when one head moves to the other, they unite in the position of the latter and will behave as one head thereafter, e.g. further movement may affect only both of them together (as one unit), never just one of them.

¹⁹ For the reanalysis affecting *mint*, see Bacskai-Atkari (2014). As far as *mert* is concerned, the change constitutes the following: interrogative operator > relative operator > complementiser. The existence of related operators in interrogatives and relatives may be familiar from several languages, e.g. German *warum* ‘why’ *weshalb* ‘wherefore’ are possible in both types of constructions. Regarding the reanalysis from relative operator into complementiser, Juhász (1991: 481–482) and Haader (2003: 542–543) identify elliptical constructions as the most likely bridging contexts. Juhász (1991: 481) provides the following example: a sentence containing a headless relative clause such as “*miért* [why-Rel] *bántottad*, (*azért* [that.for]) *haragszik*” (‘why you hurt him, he is angry for that’) could be reinterpreted as a complement clause “*mert* [because] *bántottad*, (*azért* [that.for]) *haragszik*” (‘because you hurt him, he is angry’). Note that since both headless relative clauses and complement clauses in Hungarian are allowed to have both a matrix pronominal placeholder in the matrix clause (e.g. *azért* ‘for that’), as well as no overt marker at all, the syntactic environment is rather similar for the two types constructions. Since the present paper is not devoted to the issue of changes affecting *miért/mert*, I will not elaborate on these questions any further.

The application of head adjunction results in the reverse order of the two heads, due to Kayne's Linear Correspondence Axiom (Kayne 1994); cf. also the Mirror Principle of Baker (1985, 1988). In other words, while the linear order of the two elements in their base positions is such that the higher head invariably (and naturally) precedes the lower one, head adjunction merges the original lower element to the left of the higher one and thus their order changes. (For reasons why this should be so, see also Grimshaw 1986). Hence, an original combination such as *hogymint* 'COMP than' or *hogymert* 'COMP because' was reversed to *minthogy* 'than COMP' or *merthogy* 'because COMP', respectively, as shown by the following examples:²⁰

- (7) a. semi nagob nem mondathatik: **mint** **hogh** legon
nothing greater not can.say-Pass.3.Sg. than COMP be-Subj.3.Sg.
 istenek ania
God-Dat. mother-Poss.3.Sg.
 'nothing greater can be said to be than that she is the mother of God'
 (*TihK.*, 143)
- b. Es monda, o nekýk Iwdith, **Mýnth**, **hog**, az mel'eth
and said-3Sg them they-Dat Judith as COMP that what-Acc
 zolhatéék, ýstenéének vsmeritek, Azonkeppen, az mýth
could.say-1Sg God-Poss.Dat know-2Pl likewise that what-Acc
 Ténném gondoltam, býzońýcatok megh
do-Inf.1Sg thought-1Sg prove-Imp.2Pl Prt
 'and Judith said to them: as you know that what I have been able to say is of God,
 likewise prove ye if it be of God what I intend to do prove' (*SzekK.*, 27r)

²⁰ As of 3 April 2015, the normalised part of the Old Hungarian Concordance corpus contains 3 instances of the sequence *minthogy* and 5 instances of the sequence *merthogy*, in altogether 4 codices (81947 tokens altogether). The examples in (7b) and (7c) are taken from this search; the remaining examples in (7) were searched manually based on the original spelling.

- c. Melÿ bozzosagokot frater Bernald.| bÿzon zent. nem csak
which irritations-Acc. brother Bernald indeed saint not only
 engedelmest.| de es vÿgasagost zenuediuala:| **Mert hogÿ**
obeyingly but too joyfully suffered-3.Sg.was because COMP
 bizonual uoltuolna cristusnak tekellertes tanoÿtuanÿa
indeed-Com. was-3.Sg.be-Cond. Christ-Dat. perfect student-Poss.3.Sg.
 ‘which irritations brother Bernald, indeed a saint, suffered not only obeyingly
 but also joyfully: for he was indeed a perfect student of Christ’ (*JókK.*, 20–21)
- d. azért ekkeppen alcolmas lon helheztenÿ / Harmadhoz Mÿ
therefore thus suitable was-3.Sg. place-Pass.Inf. third-All. we
 erottenc / **mert hog** meñorzagban mú kozottenc es istèn
we-Fin. because COMP heaven-Ine. we among.us and God
 kozot kozben árosnac tíztheet valya
among between walking-Dat. duty-Poss.3.Sg.Acc. has
 ‘therefore (s)he was suitable to be placed to the Trinity because (s)he has the
 task of interceding on our behalf between God and us in heaven
 (*NagyszK.*, 275)

The possible corresponding structures are shown in *Figure 8*:

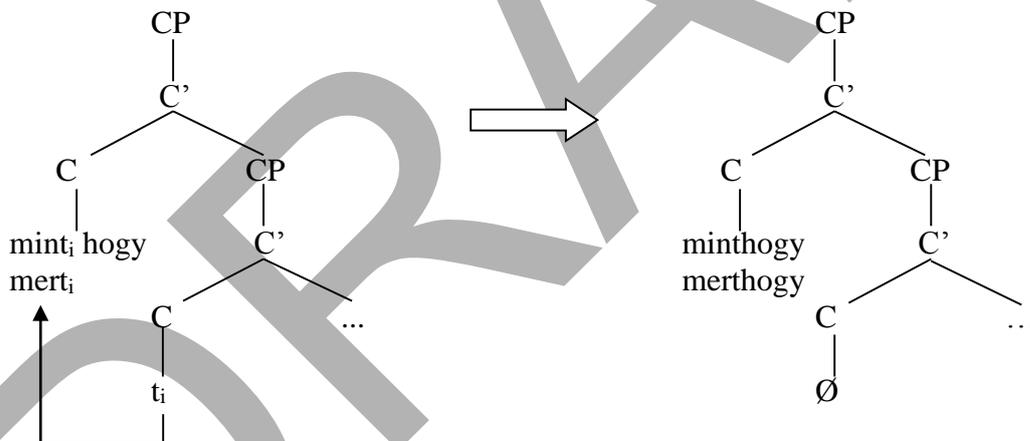


Figure 8

Let us now concentrate on the left-hand side diagram: *mint* and *mert* are base-generated in the lower C head and when they move up to adjoin *hogy* in the higher one, they are adjoined from the left and hence will appear first. Again, each such configuration had a fixed word order conforming to *Figure 8*: since typically *hogy* was originally located in the higher C (Force) head position, most combinations of merged C heads are in the X+*hogy* order – hence the configurations *minthogy* ‘than COMP’ and *merthogy* ‘because COMP’. Note that the

(original) meaning of a combination $X+hogy$ is ‘X’, just like for $hogy+X$ combinations: as a result, a given pair of $hogy+X$ and $X+hogy$ combinations, where X refers to the same complementiser, denotes two interchangeable variants.

Recall from section 4 that base-generation is preferred over movement and hence complementisers moving up from the lower C head to the higher one were ultimately reanalysed as higher C heads. This happened in the case of complex complementisers too: they started to be base-generated as complex units instead of two separate elements resulting in a combination only via movement. In this way, they became fully grammaticalised complex complementisers and their being complex is a matter of morphology and no longer of syntax. Note that since there are no single C heads base-generated in the lower C head otherwise, it would be highly circular to claim that complex complementisers are derived via movement in Modern Hungarian too; note also that if base-generation in a lower C head is allowed, there should be constructions in which the given complementiser stays in that position, which is not the case in Modern Hungarian. Hence complex C heads are fully grammaticalised; this stage is represented in the right-hand side diagram in *Figure 8*.

Since ultimately all complementisers came to be located in the higher C head, as they all started to mark the Force of the clause, the prediction is that while the fully grammaticalised complex combinations should be preserved in Modern Hungarian too, the original C+C combinations should not exist. This prediction is borne out: while *hogymint* and *hogymert* are no longer possible, Modern Hungarian still has *minthogy* and *merthogy*.

8. THE POSITION OF *HOGY*

With respect to the position of *hogy* ‘COMP’, it can be seen that the underlying order was typically of the form *hogy*+X,²¹ which is in line with the fact that *hogy* was typically located in the higher C head relatively early on anyway. Since complex complementisers surviving to the present day are invariably of the reverse order, it should not be surprising that generally combinations of the form X+*hogy* remain in the language.

There is one apparent exception, though: that of *hogy* and *ha* ‘if’, where the underlying order was *ha*+*hogy*. Considering what was said about the typical positions of *ha* and *hogy* in section 5, however, this is not in the least surprising: while *ha* was always a higher C head as early as the Old Hungarian period, *hogy* was preferably also a higher C head but could still appear in the lower C head position. Hence if they co-occurred in one CP-domain, their underlying order was naturally *ha*+*hogy*, as *ha* could not be a lower C head.

Apart from theoretical reasons, there is also independent evidence for the fact that *hahogy* reflects the underlying order: there may appear other elements in between the two CP projections (e.g. topic or focus, cf. Rizzi 1997). Hence if there is a single clause showing an intervening element between the two C heads, those C heads must be in separate projections: in other words, if there is an intervening element between *ha* and *hogy*, the string *hahogy* contains two separate C heads and hence represents the underlying order. This is indeed the case, as shown in (5b), repeated here as (8):

²¹ Note that though the underlying orders are supposed to have been present in the language earlier than the inverted ones, the examples quoted here do not show a clear temporal distinction: this is because in the late Old Hungarian and early Middle Hungarian periods the inverted orders were already possible. However, there are only few sources available from early Old Hungarian and most of these include only fragments but even the complete texts are too short to draw any conclusion with respect to whether and from when certain C+C orders were available. However, since the initial positions of the various C heads are known and grammaticalisation of complex complementisers is unidirectional, it is straightforward which combinations can be viewed as representing an underlying C+C order.

- (8) **Ha** késen **hogy** el nyugot az nap, hamar esot váry
if late COMP off set-3.Sg. the sun soon rain-Acc. expect-Imp.2.Sg.
 ‘if the sun has set late, expect rain soon’ (Cis., G3)

The left periphery of the embedded clause contains the complementisers *ha* and *hogy*; in between the two, the adverb *késen* ‘late’ can appear.²²

Since the preferred position for *hogy* was the higher C head, the tendency was that *hogy* moved up also when combined with *ha*: as a result, the reverse order (*hogyha*) was more frequent even in Old and Middle Hungarian than the underlying order (*hahogy*).²³ Again, the meaning of *hogyha* and *hahogy* is normally identical to that of *ha* ‘if’, thereby conforming to the general scheme of *hogy*+X and X+*hogy* combinations carrying the meaning of ‘X’:

- (9) a. Az én jó istenem, **ha hogy** sok ellenség, reám
the I good God-Poss.1.Sg. if COMP many enemy I-Subl.
 fegyverkezék, tolok megmente
arm they-Abl. prt-saved-3.Sg.
 ‘my good God, if many enemies armed against me, saved me from them’
 (Balassa: *Ének.*, 32)
- b. **Ha hogy** az ő keserves kin-szenvedését gyakran emlékezettel
if COMP the he bitter suffering-Poss.3.Sg.Acc. often memory-Com.
 szivedben forgatod
heart-Poss.2.Sg.Ine. rotate-2.Sg.
 ‘if you often remember his bitter sufferings in your heart’ (Csúzi: *Síp.*, 105)
- c. **Hog ha** keuelseggel kesert, te meg alazzad magadat
COMP if pride-Inst. tempts you Prt humiliate-Imp.2Sg yourself-Acc
 ‘if he tempts you with pride, humiliate yourself’ (BodK., 9v)

²² It must be mentioned that although constructions like (8) are possible, they are not typical and hence restricted in number; this is due to the fact that in Hungarian (all stages), topics and foci do not normally move to the CP-domain but to functional projections that are located below the lower CP. In other words, structures like (8) are rare because they are (and presumably were) slightly marked.

²³ As of 3 April 2015, the normalised part of the Old Hungarian Concordance corpus contains no instances of the sequence *hahogy* and 63 instances of the sequence *hogyha*, in altogether 12 codices and minor texts (179122 tokens altogether). The examples in (9c) and (9d) are taken from this search; the remaining examples in (9) were searched manually based on the original spelling.

- d. **hog ha** nē imadangatoc azon vđòbèn az ego tüznc
COMP if not worship-2Pl that-Sup time-1ne the burning fire-Dat
 kèmcèiebè èrèztètec & ki az iftèn ki tütòkèt ki
furnace-III cast-Pass.2Pl and who the God who you-Acc out
 zabadehon èn kèzèmbòl
save-Sbjv.3Sg I hand-Poss.1Sg.Ela
 ‘but if ye worship not, ye shall be cast the same hour into the midst of a burning
 fiery furnace; and who is that God that shall deliver you out of my hands?’
 (BécsiK., 127)

The structural changes affecting *hogy* in combinations with *ha* are summarised in Figure 9:

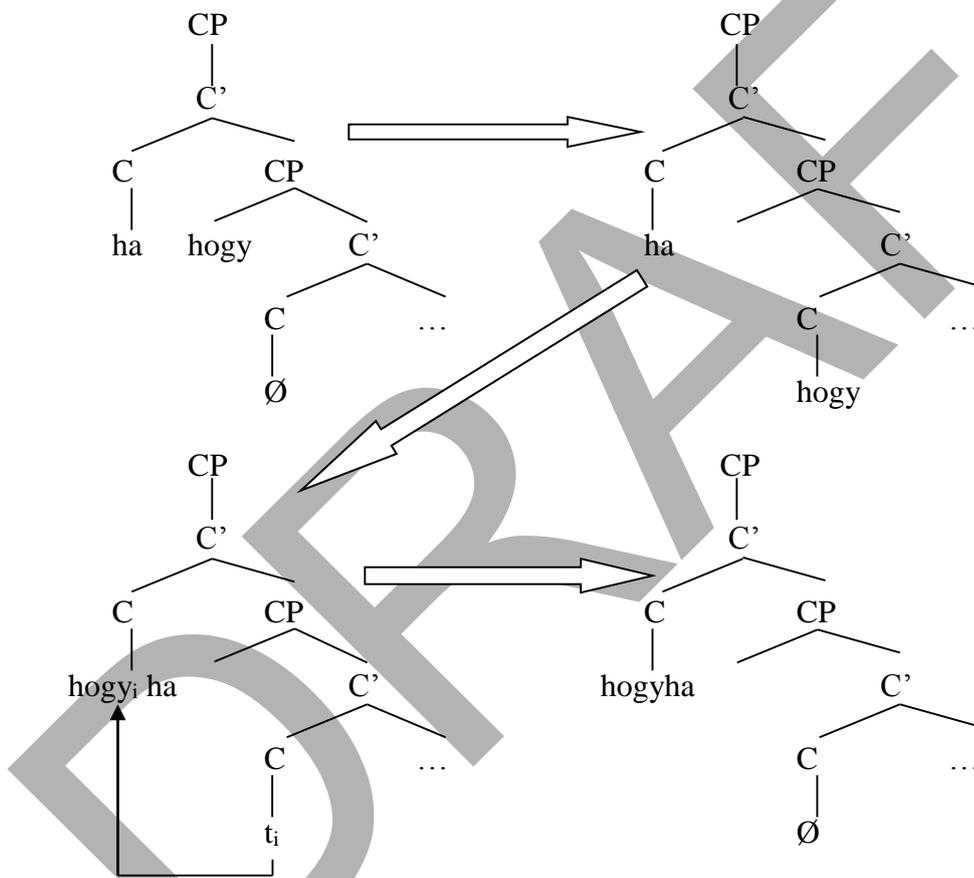


Figure 9

The individual stages are in line with the ones given in Figure 7 and Figure 8: the only difference is that *hogy* starts from the lower C head position.

9. RELATIVE CLAUSES

Apart from the complex complementiser combinations mentioned so far, *hogy* ‘COMP’ took part in other combinations in the left periphery: ordinary relative clauses could also contain

the sequence of *hogy*+ a relative operator both in Old and in Middle Hungarian, although it was enough for a relative clause to be introduced by the operator (cf. Juhász 1992: 792; Galambos 1907: 14–18; Bacskai-Atkari 2011: 112–113). Examples are shown in (10): (10a) demonstrates a relative clause with a pronominal head (*olyat* ‘such’), and (10b) demonstrated a relative clause with a lexical head (the nominal expression *hamis tanúságot* ‘false witness’).

- (10) a. olÿaat tezoek rajtad **hog** **kÿtol** felz
such-Acc. do-1.Sg. you-Sup. COMP who-Abl. fear-2.Sg.
 ‘I will do such a thing to you that you are afraid of’ (*SándK.*, 28)
- b. Az papy ffeyedelmek kedyglen, es a" venek, es mynd az
the priestly princes in.turn and the elders and all the
tellyes tanacz, kerefnek vala hamyff tanofagot Jefus ellen,
full council seek-3.Pl. was-3.Sg. false witness-Acc. Jesus against
hog mywel hwet halarra adhatnaak; es nem
COMP what-Instr. he-Acc. death-Subl. can.give-Cond.3.Pl. and not
lelenek, myert fok hamyfl tanok tamattanak vona
found-3.Pl. because many false witnesses arose-3.Pl. be-Cond.3.Sg.
 ‘Now the chief priests, and elders, and all the council, sought false witness against Jesus, with which they could put him to death; But found none: yea, though many false witnesses came.’ (*JordK.*, 443)

As can be seen, the complementiser *hogy* is followed by the operator *kitől* ‘who-Abl.’ in (10a) and *mivel* ‘what-Instr.’ in (10b). Such combinations involve *hogy* and the operator *ki* ‘who’ or *mi* ‘what’, which could be overtly marked for case (that is, any case other than the nominative). Note that *hogy* was not exceptional with respect to its ability to combine with operators in relative clauses: there are several examples with the complementiser *ha* ‘if’ as well (without such clauses being conditional, however). Consider:

- (11) a. kÿ tegod zereth. az nem epedh: **ha** **kÿ** keserg akkor wÿgad
who you-Acc. loves that not longs if who moans then rejoices
 ‘those who love you, do not long: those who moan, then rejoice’
 (*CzechK.*, 51–52)

- b. Bizoń bizō mōdom tunètec **ha mit** keèndetec [sic!] èn
indeed indeed say-1.Sg. you-Dat. if what ask-Fut.2.Pl. I
 at'amtol èn nèuembè / agga tunètec
father-Poss.1.Sg.Abl. I name-Poss.1.Sg.Ill. gives you-Dat.
 ‘Verily, verily, I say unto you, Whatsoever ye shall ask the Father in my name,
 he will give it you.’ (MünchK., 103ra)

The left periphery of the subclause in (10a) – containing the complementiser *hogy* and the relative operator *kitól* – is shown in *Figure 10*:

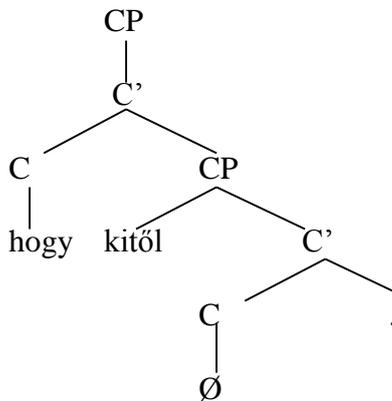


Figure 10

As can be expected, *hogy* occupies the higher (Force) C head position, while the operator is located in the specifier of the lower CP, conforming to the general pattern of the CP-domain outlined in section 1; the same would be true for *ha* + operator combinations. This configuration is similar to the precursor of C+C combinations (see *Figure 7*), where the future complementisers *mint* and *mert* were still operators and hence located in the lower [Spec,CP]. Since relative operators did not develop into C heads, the reverse order of *hogy* and the relative operators was not possible: accordingly, no such examples are attested. This should not be surprising: recall from section 4 that the reanalysis of operators into C heads is possible only if they lose their functions that they used to have in the TP-domain. In the case of *ki* and *mi* this condition was clearly not met as these elements were base-generated as arguments of the lexical verb and hence their base-generation as C heads would have severely affected the argument structure of the clause. It logically follows that operators retaining their original

person and number features will not be reanalysed as C heads. Note that this is a restriction that holds in Hungarian but, as was mentioned earlier, there are languages that may permit complementiser agreement – for instance, as described by Willis (2007), the reanalysis of an expletive subject in Welsh to an affirmative complementiser does not involve the loss of agreement features but it necessarily involves the loss of the particle *a*, C heads by definition not being allowed to be phrase-sized (see Willis 2007: 450–454).

10. FUNCTIONS OF *HOGY*

Since *hogy* ‘COMP’ had the ability to appear in a large variety of constructions, the issue of its functions must also be addressed.

With respect to comparatives, the examples discussed so far might suggest that *hogy* in these constructions was subsidiary to *mint* ‘than/as’. This is, however, not the case: the original comparative complementiser was actually *hogy* (cf. the discussion at the end of section 8), and *mint* started to occur in Old Hungarian: first as an operator and subsequently it was reanalysed as a lower C head, following the mechanism of the relative cycle (cf. Bacskai-Atkari 2011, 2014). As *mint* started to be interpreted as the element responsible for introducing comparative Force in the subclause, *hogy* gradually lost the same function.

On the other hand, with the loss of specific functions, *hogy* became the general marker of subordination. This was accompanied by functional extension: *hogy* appeared in other clauses functioning as a general subordinating C head. There are two main pieces of evidence for this that will be familiar from the previous sections. First, *hogy* could appear in a wide range of clauses, such as conditionals, clauses of reason, relative clauses, or embedded interrogatives. Second, the meaning of a combination *hogy*+X or X+*hogy* did not (initially) differ from the meaning of X, which is a clear indication of *hogy* being a marker of a functional syntactic property.

Later, as other complementisers started to consistently mark subordination besides their specific functions, *hogy* was no longer used as a general subordination marker and hence was no longer combined with other elements in the way it used to be. This naturally contributed to the disappearance of *hogy* in relative clauses but it did not affect already grammaticalised complex complementisers as they were fossilised syntactic units. Marking subordination is not independent of the exact syntactic position of a given C head: it is associated with the higher (Force) C head and hence if an element is interpreted as a subordination marker, it is expected to appear in that position. In other words, the fact that other complementisers also started to be markers of subordination implies that they started to be base-generated as higher C heads, which naturally prohibits the base-generation of *hogy* as an independent head in the very same position; hence *hogy* naturally disappeared from these constructions.

11. CONCLUSION

The aim of this article was to investigate the issue of the Hungarian declarative complementiser *hogy* 'COMP'. It was shown that *hogy* developed via the relative cycle from an original operator into a lower C complementiser head, which was later reinterpreted as a higher C head, in the same way as English *that* developed. I also demonstrated that *hogy* originally had other, more specific functions too but was gradually interpreted as a general marker of subordination and hence came to be used in a wide range of constructions.

Since it frequently appeared together with other complementisers, it also took part in the formation of complex complementiser units. With the development of all the other complementisers into higher C heads, *hogy* was no longer necessary to mark subordination separately – hence only those complex complementisers remain in the language that were fully grammaticalised into a single C head.

LIST OF TEXTUAL SOURCES

- AporK. = Apor-kódex [Apor Codex]. End of the 15th century – beginning of the 16th century.
- Balassa: Ének. = Balassa Bálint énekei [The songs of Bálint Balassa]. Levoča 1693.
- BécsiK. = Bécsi kódex [Vienna Codex]. Middle of the 15th century.
- BodK. = Bod-kódex [Bod Codex]. First half of the 16th century.
- Cis. = Cisio. Cluj-Napoca 1592.
- Csúzi: Síp. = Csúzi Zsigmond. Zengedező sípszó [Zsigmond Csúzi: Praising Whistle].
Pressburg 1723.
- CzechK. = Czech-kódex [Czech Codex]. 1513.
- GuaryK. = Guary-kódex [Guary Codex]. Before 1508.
- JókK. = Jókai-kódex [Jókai Codex]. 14th–15th century.
- JordK = Jordánszky-kódex [Jordánszky Codex]. 1516, 1519.
- KazK. = Kazinczy-kódex [Kazinczy Codex]. 1526–1541.
- KTSz. = Königsbergi töredék és szalagjai [Königsberg Fragment and its Ribbons]. Middle or second half of the 14th century.
- LázK. = Lázár-kódex [Lázár Codex]. After 1525.
- MünchK = Münchener kódex [Munich Codex]. 1466.
- NagyszK. = Nagyszombati kódex [Nagyszombat Codex]. 1512, 1513.
- SándK. = Sándor-kódex [Sándor Codex]. The first quarter of the 16th century.
- SzekK. = Székelyudvarhelyi kódex [Székelyudvarhely Codex]. 1526–1528.
- TihK. = Tihanyi-kódex [Tihanyi Codex]. 1532.

REFERENCES

- AXEL, KATRIN, 2009. 'Die Entstehung des dass-Satzes – ein neues Szenario', in Veronika Ehrich, Christian Fortmann, Ingo Reich & Marga Reis (eds.), *Koordination und Subordination im Deutschen*, Hamburg: Buske, 21–41.
- BACSKAI-ATKARI, JULIA, 2011. 'A komparatív operátor esete a mondatbevezetővel:

- Szintaktikai változások a magyar hasonlító mellékmondatokban', in Katalin É. Kiss & Attila Hegedűs (eds.), *Nyelvelmélet és diakrónia*, Budapest & Piliscsaba: Szent István Társulat, 103–119.
- BACSKAI-ATKARI, JULIA, 2012. 'Reducing Attributive Comparative Deletion', *The Even Yearbook* 10, 1–25.
- BACSKAI-ATKARI, JULIA, 2014a. 'A véges alárendelő összetett mondatok története', in Katalin É. Kiss (ed.), *Magyar generatív történeti mondattan*, Budapest: Akadémiai Kiadó, 239–279.
- BACSKAI-ATKARI, JULIA, 2014b. 'Cyclical Change in Hungarian Comparatives', *Diachronica* 31.4, 465–505.
- 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.
- BAYER, JOSEF, 1984. 'Comp in Bavarian Syntax', *Linguistic Review* 3, 209–274.
- BAYER, JOSEF & ELLEN BRANDNER, 2008. 'On Wh-Head-Movement and the Doubly-Filled-Comp Filter', in Charles B. Chang & Hannah J. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, Somerville, MA: Cascadilla Proceedings Project, 87–95.
- CHOMSKY, NOAM, 1977. 'On WH-movement', in Peter W. Culicover et al. (eds.), *Formal Syntax*, New York: Academic Press, 71–132.
- CHOMSKY, NOAM, 1995. *The Minimalist Program*, Cambridge, MA: MIT Press.
- É. KISS, KATALIN, 2002. *The Syntax of Hungarian*. Cambridge: Cambridge University Press.
- EGEDI, BARBARA & ESZTER SIMON, 2012. *Gradual Expansion in the Use of the Definite Article. Checking a Theory against the Old Hungarian Corpus*. Talk delivered to: Exploring Ancient Languages through Corpora, Oslo, 14–17 June 2012. www.nytud.hu/oszt/elmnyelv/mgtsz/2013/122_egedi-simon_oslo_prezi.pdf
- GALAMBOS, DEZSŐ, 1907. *Tanulmányok a magyar relatívum mondattanáról*, 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.
- GRIMSHAW, JANE, 1986. 'A Morphosyntactic Explanation for the Mirror Principle', *Linguistic Inquiry* 17.4, 745–749.
- G. VARGA, GYÖRGYI, 1992. 'A névmások', in Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana II/1.: A kései ómagyar kor: Morfematika*, Budapest: Akadémiai Kiadó, 455–569.
- HAADER, LEA, 1991. 'Az alárendelő mondatok: Az alanyi, állítmányi, tárgyi és határozói mellékmondatok', in Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana I.: A korai ómagyar kor és előzményei*, Budapest: Akadémiai Kiadó, 728–741.
- HAADER, LEA, 1995. 'Az alárendelő mondatok: Az alanyi, állítmányi, tárgyi és határozói mellékmondatok', in Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana II/2.: A kései ómagyar kor: Mondattan. Szöveggrammatika*, Budapest: Akadémiai Kiadó, 506–665.
- HAADER, LEA, 2003. 'Az ómagyar kor: Mondattörténet: Az összetett mondat', in Jenő Kiss & Ferenc Pusztai (eds.), *Magyar nyelvtörténet*, Budapest: Osiris Kiadó, 500–560.
- JUHÁSZ, DEZSŐ, 1991. 'A kötőszók', in Loránd Benkő (ed.), *A magyar nyelv történeti*

- nyelvtana I.: A korai ómagyar kor és előzményei*, Budapest: Akadémiai Kiadó, 476–500.
- JUHÁSZ, DEZSŐ, 1992. ‘A kötőszók’, in Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana II/1.: A kései ómagyar kor: Morfematika*, Budapest: Akadémiai Kiadó, 772–814.
- KÁNTOR, GERGELY, 2008. ‘On Hungarian Relative Operators’, *The Even Yearbook* 8, 1–12.
- KAYNE, RICHARD, 1994. *The Antisymmetry of Syntax*, Cambridge, MA: MIT.
- KENESEI, ISTVÁN, 1992. ‘On Hungarian Complementizers’, *Approaches to Hungarian* 4, 37–50.
- KENNEDY, CHRISTOPHER & JASON MERCHANT, 2000. ‘Attributive Comparative Deletion’, *Natural Language & Linguistic Theory* 18, 89–146.
- RIZZI, LUIGI, 1997. ‘The Fine Structure of the Left Periphery’, in Liliane Haegeman (ed.), *Elements of Grammar*, Dordrecht: Kluwer, 281–337.
- ROBERTS, IAN, 2005. *Principles and Parameters in a VSO Language*, Oxford, Oxford University Press.
- ROBERTS, IAN & ANNA ROUSSOU, 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization*, Cambridge: Cambridge University Press.
- SIPOS, PÁL, 1991. ‘A névmások’, in Loránd Benkő (ed.), *A magyar nyelv történeti nyelvtana I.: A korai ómagyar kor és előzményei*, Budapest: Akadémiai Kiadó, 353–400.
- WILLIS, DAVID, 2007. ‘Specifier-to-Head Reanalyses in the Complementizer Domain: Evidence from Welsh’, *Transactions of the Philological Society* 105:3, 432–480.