From parataxis to finite subordination in the Ugric languages and beyond

1. Introduction

This paper documents an evolutionary path from parataxis via the correlative construction to finite subordination in the Ugric languages, and it also points out parallels in the the Indo-European language family.

Finite subordination was traditionally assumed to have evolved from parataxis, i.e., the juxtaposition of logically linked simple sentences – see e.g. Delbrück (1900: 411). Delbrück and many of his contemporaries argued that early languages such as Proto-Indo-European, in which no finite subordination can be reconstructed, only used simple sentences, and the emergence of finite subordination was an advancement in their evolution. This assumption, however, lost credibility with the advent of generative linguistic theory, according to which human languages at all times are manifestations of the same biologically determined universal grammar. It has been found that early languages only lacked *finite* subordination; they used non-finite subordination instead, as is typical of SOV languages (Lehmann 1980; É. Kiss 2023a). Naturally, the presence of non-finite subordination in early languages does not exclude the evolution of paratactic constructions into finite subordination at the same time, and this possibility (e.g., the change of *I think that. The Earth is round* into *I think that the Earth is round*) has indeed been raised not only in connection with complement clauses (Davidson 1968; Hopper and Traugott 2003: 190–194; Roberts and Roussou 2003: 116–121; van Gelderen 2004: 81; Heine and Kuteva 2007: 241; Narrog and Heine 2017: 8–27), but also in connection with relative clauses introduced by a relativizer grammaticalized from a demonstrative or a *wh*-word (O'Neil 1977; Hopper and Traugott 2003: 196–204; Heine and Kuteva 2007: 224–229; Narrog and Heine 2017: 8–27). At the same time, Harris and Campbell (1995:282), Roberts (2007), and Walkden (2018), among others, have questioned the hypothesized developmental paths claiming that present-day languages, including languages with a well-documented history, display no uncontroversial cases of parataxis to finite subordination evolution. Hence in view of the uniformitarian principle, according to which the processes that were active in the evolution of human languages are those that we can observe today, it is unlikely that the attested finite subordination patterns have developed from paratactic constructions. Harris and Campbell (1995:300) assume that finite subordinate clauses could have developed from non-finite verb forms by the use of the same phrase structure rules that were available for the construction of simple sentences – even though present-day languages provide no examples of this kind of change, either.

Recently, it has convincingly been argued by Axel-Tober (2017) that Germanic *that*-type complementizers were originally relative pronouns introducing the relative clause in the so-called explicative construction (e.g. *I think* [*it* [*that the Earth is round*]]). Weiss (2020) has extended this assumption to other languages, claiming that the grammaticalization of complementizers is only possible from a hypotactic source structure, often a structure containing a relative clause headed by a light noun or a demonstrative pronoun. Axel-Tober's and Weiss's assumptions converge with the widely accepted hypothesis that Proto-Indo-European had finite subordinate clauses only in the form of relative clauses, and the oldest type of finite relative construction is the correlative structure involving a headless relative clause coindexed with a main clause anaphoric element (Kiparsky 1995; Ram-Prasad 2022). Weiss raises the possibility that this structure may have been the first type of subordination to arise crosslinguistically. As for the source of relative clauses, "there is no reliable evidence for the parataxis-to-hypotaxis hypothesis .... The standard opinion [the purported development from parataxis to hypotaxis] has proven wrong and must be abandoned" (Weiss 2020:50).

The present paper claims based on evidence from the Ugric branch of the Uralic language family that the correlative construction and the explicative construction identified as the

sources of complex sentences with a finite embedded clause are intermediate stages in the development from parataxis to finite subordination. The argumentation is structured as follows: Section 2 presents the types of propositional complements and propositional adjuncts used in the Ob-Ugric languages (Khanty and Mansi), claiming that parataxis is not an earmark of less advanced grammar but fills a special role in their strictly SOV syntax. Section 3 discusses the first finite subordinate clause type emerging in the Ob-Ugric languages: the correlative construction, pointing out its paratactic origin. Section 4 discusses how the correlative construction evolved into finite relativization in Hungarian, a Ugric language that has drifted from SOV to SVO. Section 5 shows the emergence of Hungarian finite complement clauses employing the relative pronoun of a correlative construction as their complementizer. Section 6 is a summary.

2 Propositional complements in SOV Ob-Ugric: Non-finite subordination and parataxis

Khanty and Mansi, the Ugric languages spoken along the river Ob, preserve basic properties of the syntax attributed to Proto-Uralic by comparative Uralic linguistics (Ravila 1960; Jahnhunen 1981; Aikio 2019). Khanty and Mansi texts recorded before the 1960s from monolingual or Khanty/Mansi-dominant bilingual speakers display a strict SOV word order. As is typical in strict SOV languages, propositional complements, propositional adjuncts, and propositional modifiers are expressed by non-finite constructions – as illustrated by a Khanty complement clause (1a), adjunct clause (1b), and a prenominal subject gap relative clause (1d):

(1)a. [**pro tu:ta xoj-na ji-te:-mən**] śa:t-l
(we) here who-LOC come-PRSPTCP-1DU be.heard-PRS.3SG
'It is heard that we two are visited by someone.'
(Northern Khanty, Nikolaeva 1999: 49)

b. ma$_i$ [**pro$_i$ kanək wåńť-ťaɣə mən-t-am-nə**] wičipə ənəł qynt wə-ł-əm.
I berry gather-INF go-PRSPTCP-1SG-LOC always big basket take-PRS.1SG
'When I go to gather berries, I always take a big basket [with me].'
(Northern Khanty, Nikolaeva 1999: 48)

c. [**Ø$_i$ kŏł-nȧm qołəɣtə-łəɣ**] ńewrəm-əm$_i$ pətȧn mȧntem ȧtəm-ɣə pit.
word-APR listen-PTCP.NEG child-1SG because.of I-DAT bad-TRA become.PST.3SG
'I felt bad because of my child who is not listening to word.'
(Eastern Khanty, Csepregi 2023: (32))

The non-finite verbal projections have various features associated with finiteness. They can have an independent subject (represented by a 1st person dual pro in (1a)), and as shown by (1a,b), their subject can elicit agreement on the non-finite verb. They can involve negation (1c). They can express aspectual differences, i.e., simultaneity (1b) and anteriority (2a,b), by choosing the appropriate participial or converbial suffix in relative clauses (2a), and by supplying the clause with an appropriate case suffix or postposition in adverbial clauses (2b).

(2)a. pro [**ma Ø$_i$ mən-m-am**] lek-i$_i$ jüw-a!
(you) I go-PSTPTCP-1SG road-ABL arrive-IMP.2SG
'Arrive on the road on which I went!'
(Csepregi, fieldwork)

b. ma$_i$ [**pro$_i$ quj-a mən-t-am-a**] jăqə əj nŏrɣə måjəł-taɣə jăŋqił-ł-əm.

I man-LAT go-PRSPTC-1SG-LAT home continually visit-INF go-PRS-1SG
'After I have got married, I will continually go home for a visit.'
(Csepregi, fieldwork)

The non-finite verb can also be marked for conditional mood:

(3) [**t'umint süj-əł lirti pit-t-ał-ka**] nüŋ mən-a panə kemə łiwt-a
such noise-3SG arise- PRSPTC-3SG-COND you go-IMP.2SG and out exit- IMP.2SG
'If that noise arises, you should go out'
(Csepregi 2011: 69)

Despite their rich functional domain, non-finite clauses cannot have either independent tense or independent force. Propositional complements with independent tense and/or independent force, primarily complements of reporting predicates, need to be expressed as independent sentences:

(4) pɯːrəs iːki jɐːstə-ɬ: [**pro qoɬnə βɛt-tɐːə moːst-əɬ?**]
old man say-PRS.3SG how kill-INF need-PRS.3SG
'The old man says: How shall he be killed?'
(Eastern Khanty, Paasonen 1901, OUDB 1313)

b. pro āśe-n kitiɣl-awe: [**piɣkʷe, manəriɣ Íuńś-eɣn?**]
father-LAT ask-PASS.3SG boy.DIM why cry-2SG
'It is asked by the father: Little boy, why are you crying?'
(Mansi, Kannisto 1951, OUDB 1143)

As shown by Szeverényi and Sipőcz (2019) in their study of reportative constructions in Mansi and the related Nganasan language, propositions are reported paratactically also in current spontaneous, everyday conversations in these SOV languages.

The distribution of non-finite subordination and parataxis in the strictly SOV Ob-Ugric languages (as well as in Nganasan) indicates that parataxis is not the manifestation of less advanced, less sophisticated syntax; it fills a special niche in SOV languages with non-finite subordination.

3 Emergence of the correlative construction

3.1 Paratactic origin in Ob-Ugric

The only sentence type in traditional Ob-Ugric texts that appears to the present reader to involve two finite clauses linked in an asymmetric relation, i.e., to represent finite subordination, is the so-called correlative construction. The correlative construction, best known from the Indo-Aryan languages (Srivastav 1991; Dayal 1996; Bhatt 2003), has the following characteristic properties (Lipták 2009): It consists of a free relative clause, and a subsequent main clause containing an expression (a DP, PP, or an adverb) that is anaphorically related to the relative pronoun of the initial clause – as illustrated by the Hungarian example in (5a). Locality and binding tests indicate that the relative clause is adjoined to the main clause (Lipták 2012). The relative clause can contain more than one relative expression with an anaphoric correlate in the main clause (5b). The main clause correlate of the relative phrase is often a demonstrative phrase, or an overt or dropped personal pronoun. The NP modified by the relative clause can be spelled out either in the relative pronoun (6a), or in the demonstrative phrase (6b), or in both (6c).

(5)a. **Aki** nem dolgoz-ik, **az** ne is e-gy-ék.
who not work-3SG that not either eat-SBJV-3SG
'He who does not work, neither shall he eat.'
(Hungarian)

b. **Ki** **mint** vet, [**az**] **úgy** arat.
who how sow.3SG that so reap.3SG
'As one sows, so he reaps.'
(Hungarian)

(6)a. **Amelyik** **látogató** az interneten vesz jegy-et, **an-nak** nem kell
which visitor the internet-on buy.3SG ticket-ACC that-DAT not need.3SG
sor-ba áll-ni-a a pénztár-nál.
line-in stand-INF-3SG the ticket_office-at
'The visitor who buys a ticket online need not stand in line at the ticket office.'

b. **Aki** az interneten vesz jegy-et, **an-nak** **a** **látogató-nak** nem kell
who the internet-on buy.3SG ticket-ACC that-DAT the visitor- DAT not need.3SG
sor-ba áll-ni-a a pénztár-nál.
line-in stand-INF-3SG the ticket_office-at

c. **Amelyik** **látogató** az internet-en vesz jegy-et, **an-nak** **a** **látogató-nak**
which visitor the internet-on buy.3SG ticket-ACC that-DAT the visitor- DAT
nem kell sor-ba áll-ni-a a pénztár-nál.
not need.3SG line-in stand-INF-3SG the ticket_office-at
(Hungarian)

The correlative clause has maximalizing semantics (Grosu and Landman 1998; Belyaev and Haug 2020), which is in most cases manifested in a free choice reading with a universal force, rendered in English by a *whoever, whatever, whenever*-type relative pronoun. At the same time pragmatics, e.g., a once-only predicate, can restrict maximality to a unique referent, e.g.:

(7) **Aki** első-nek érkez-ett, **az** be-kapcsol-t-a a fűtés-t.
who first-DAT arrive-PST.3SG that on-turn-PST.3SG the heating-ACC
'He who arrived first turned on the heating.'
(Hungarian)

The Khanty and Mansi correlative constructions appear to satisfy the syntactic criteria listed above. They consist of a headless relative clause and a subsequent main clause containing a demonstrative phrase anaphorically related to the relative pronoun.

(8)a. [**möγöl'i** mä-nä mas-wəl] pro **t'u** məjiγilə-γäs.
what I-LOC need-PRS.3SG that give-PST.3SG
'**What** is needed for me, **that** he gave.'
(Eastern Khanty, Gulya 1966: 86)

b. [**kumlʲə** tʲe uj təːn pæl kart-au] **ækʷ** **tʲe** **kajtəl** næu koɒ̯t-æn
how this bear sinew tighten-PASS.3SG same_way you hands-2SG
ləil-æn pær kart-ɔŋkʷ-ət.
feet-2SG shrivel-PASS.IMP-3PL
'**How** this bear's sinew is tightened, **so** shall your hands and feet be shriveled.'

(Mansi, Munkácsi 1896, OUDB 1419)

The relative clause can contain more than one relative expression:

(9) [**qŏjaɣi** ńuʌ-əʌ **qŏʌə** körɣ-əʌ] panə **pro ťu** **tåɣi-nə** ne ťi
who arrow-3SG where fall-PST.3SG and that place-LOC woman so
wəj-itəɣ.
take-IMP.2PL
'**Where** the arrow of **each of you** falls, **there you** shall take a wife.'
(Eastern Khanty tale, cited by Csepregi 2011:34)

Like in the Hungarian (6a), the relative phrase can contain a lexical NP:

(10) [puːt βərrənt-əttə tɐːə-ji **məβ** **ɐːrit** **ɬiːtot** qɯːtʲ-ɬʲ]
pot make-PRSPTCP place-ABL how much food remain-PRS.3SG
pro$_{subj}$ qołtɐːəł mən-t-ɐ łɐːt-nə joːt-ɐ **pro$_{obj}$** tuː**-iłɐ**
tomorrow leave-PRSPTCP-2SG time-at with-2SG take-IMP.**OBJ**.2SG
'**What food** is left in the cooking spot, take **them** with you when you leave tomorrow!'
(Eastern Khanty, Paasonen 1901, OUDB 1313)

The Ugric languages have both subject and object agreement and both subject and object pro-drop, and the main clause correlate of the relative phrase can be a pro cross-referenced on the verb, as shown in (10) and (11):

(11) [**manər** al-as-lum] jaltapt-**eln**!
what kill-PST-SG.1SG revive-IMP.**OBJ**.2SG
'**What** I killed, revive **it**! '
(Mansi, Černecov 1935, OUDB 1234)

Temporal and locative relative adverbs can also have implicit main clause correlates. Dékány et al. (2020) analyze example (12b) (their example (2)) analogously to (12a), as a correlative construction with an implicite *there*:

(12)a. [**χotal'** nomt-e pat-i] **tuw** wos min-i.
where thought-3SG fall-PRS.3SG there PRT go-PRS.3SG
'**Where** his thought tends, **there** he goes.'
(Mansi, Kannisto and Liimola 1951, cited by Riese 2001:72)

b. [Pupi **qŏt** ŏjəɣtə-s-təɣ], jəm uɬəm wär-s-əɣən pupi-nat.
bear where find-PST-OBJ.3SG good dream do-PST-3DU bear-INS/COM
'**Where** he found the bear, [**there**] they said goodbye [to each other] with the bear.'
(Eastern Khanty, Paasonen 1901, OUDB 1315)

Regarding semantics, however, these Ob-Ugric constructions do not match the canonical correlative construction; the relative clause does not have the expected universal or free choice reading. In the early Ob-Ugric texts, its reading is episodic – as is obvious also without a wider context in the case of (8b), (10), (12b), or (13).

(13) [**xot** ti pox xaln tara min-i] at **ti** suitiɣl-i!

who this side along go_across-PRS.3SG not this sound-PRS.3SG
'**Who** goes on this side across, **this** doesn't make any noise.'
(Mansi, Černecov 1935, OUDB 1235)

The first known Ob-Ugric examples with a clear universal interpretation, e.g. (14a), a proverb translated from Russian, and (14b), were recorded in 1956-57. In (14b), the universal reading is supported by a universal morpheme.

(14)a. [**koji** əntə ropiltə-wəl, **pro** əntə li-wəl.
who not work-PRS.3SG not eat-PRS.3SG
'**Who** does not work, **he** does not eat.'
(Eastern Khanty, Gulya 1966: 86)

b. [**kol-əpa** kit-l-im] **toɣ-əpa** mən-äti.
where-all send-PRS-SG.1SG there-all go-IMP.SG.SG2
'**Wherever** I send him, **there** he shall go.'
(Eastern Khanty, Gulya 1966: 142)

There are also more compelling reasons to doubt that the construction in question functioned as a correlative construction in traditional Ob-Ugric. First, the Ob-Ugric languages had no relative clauses introduced by a relative pronoun at the time when the early texts were recorded. Relative pronouns appeared only recently in the language of the Russian-Khanty, Russian-Mansi bilingual generations. In traditional Khanty and Mansi texts, we only find prenominal participial gap relatives, as illustrated in (1c), (2a), and (15):

(15) [Øi ńa:vre:m-l-al wo:s-na man-əm] purəś ike:-ti
child-PL-3SG city-LOC go-PSTPTCP old man-PL
'the old men whose children went to the city'
(Northern Khanty, Nikolaeva 1999:77)

The pronouns translated todays as relative pronouns were in all probability indeterminate pronouns (Kratzer and Shimoyama 2002), i.e., pronouns introducing sets of alternatives bound by a covert (mostly interrogative or existential) propositional operator. Nikolaeva's Khanty grammar (Nikolaeva 1999:18) assigns interrogative and indefinite readings to them. In DEWOS, the large Khanty dictionary (Steinitz 1967–1991), too, several dialectal versions of the Khanty *wh*-pronouns in (8)–(14) occur with interrogative and indefinite readings. In the above sentences the interrogative readings do not result in coherent interpretations; the *wh*-pronouns must function, or must have functioned originally, as indefinites. For example, (13) must have meant for the speaker who originally uttered it 'Someone goes on this side across; (s)he doesn't make any noise.'

Another reason to assume that the *wh*-pronoun in the initial clause of the apparent correlative construction must have been interpreted indefinitely is that a conjunctive particle may appear between the two clauses. In Khanty, its appearance is sporadic (see example (9)), but in Pelym Mansi *when* clauses it appears regularly – see (16a,b). In the correlative rendering of these sentences, the conjunction has to be ignored and an implicit demonstrative has to be assumed, but if the *wh*-pronoun is interpreted as an indefinite, the coordinating conjunction has a natural role.

(16)a. [pær **kʷæʃ** aŋkeə̯t-i **joŋ** kulʲ jeə̯jənt-i
back when look-PRS.3SG and devil come-PRS.3SG

‘He looks back sometime, and the devil is coming.’
(rather than ‘When he looks back, [then] the devil is coming.’)
(Pelym Mansi, Kannisto 1955, OUDB 1268)

b. [**kʷæʃ** min-iː-ɣ joɒ̯ axtəʃ nʲal oɒ̯tʲəm
when go-PRS-3DU and stone promontory NEG.EXIST
‘They go there, and the stone promontory does not exist.’
(rather than ‘When they go there, [then] the stone promontory does not exist.’
(Pelym Mansi, Kannisto 1951,1278)

Notice that in (16b), the equivalent of *when* is not in clause-initial position as would be expected if it were a relative pronoun in the specifier of a CP projection. (Neither is the *wh*-pronoun in initial position in (10) and (12b)).

These facts (the episodic reading of the alleged correlative clause, the lack of *wh*-relatives at the time when the above correlative constructions were recorded, the occasional presence of a conjunction between the two clauses, and the *in situ* position of the wh-pronoun) suggest that the apparent correlative constructions of traditional Ob-Ugric texts are not subordinate structures; they are paratactically linked sentences, or, occasionally, two sentences linked by a conjunction. The pronouns of the initial clauses are indeterminate pronouns interpreted indefinitely. At least this is how they were understood when these texts were recorded.

Their interpretation, however, may have changed by now. As shown by Dékány et al. (2020), recently finite adnominal relative clauses introduced by a relative pronoun have appeared in Eastern Khanty. Relative pronouns are formally non-distinct from interrogative/indeterminate pronouns. The analysis of the initial clause of the construction in question as a relative clause is presumably supported by the correlative interpretation of similar constructions in Russian, the dominant language of most present-day Khanty speakers.

3.2 Relics of paratactic origin in Old Hungarian correlatives

Hungarian, the third Uralic language, provides further evidence testifying that the pronoun in the initial clause of correlative constructions was originally an indefinite pronoun.

Hungarian has evolved into an SVO language with dominantly finite subordination since the time it parted from its Ugric sister languages around 500 BC.[1] Old Hungarian (documented in coherent texts from 1195), however, still displays several constructions similar to those preserved in present-day Khanty and Mansi, even if they are gradually losing ground to structures typical of head-initial syntax. Thus non-finite constructions still represent a large – though decreasing – proportion of Old Hungarian subordinate clauses. Correlative constructions are also common, and their initial clauses still include a universally interpreted indeterminate pronoun.

Hungarian indeterminate pronouns have been analyzed by Bende-Farkas (2014). As she demonstrates with Old Hungarian examples, an indeterminate pronoun, e.g. *ki* ‘who’, can be bound by an invisible interrogative operator (17a), by existential closure or by the conditional *ha* ‘if’ yielding a free choice reading (17b), by negation (17c), or by a universal quantifier (17d).

(17) a. **Ki-c** oʒvc. miv vog-muc.
who-PL those we be-1PL
‘**Who** are those? It is us.’

[1] Recent archeological and archeogenetic research suggests that the split of Hungarians from their Ob-Ugric sisters could have happened much later, in the 7th century AD (Német and Vígh 2019), or just a few decades before the arrival of Hungarians in the Carpathian Basin in 895 (Türk 2023).

(Old Hungarian, *Halotti beszéd,* 1195)

b. De ha **kÿ** kerde-ne mÿ lee-gÿ-en az eredet zerent wal-o
but if who ask-COND.3SG what be-SBJV-3SG the origin according.to be-PTCP
ÿgassagh .... wgÿ mond zent anselmws doctor
truth so says Saint Anselmus doctor
'But should **somebody** ask what the original truth is ... doctor Saint Anselmus says... '
(OH, Érsekújvári Codex, 1529: 289r)

c. Az ÿo lelkew ember-nek kedeeg nagÿ-ob erdem-ót zerez vel-e
the good natured man-DAT CONJ bigg-er merit-ACC obtain.3SG INS-3SG
chak **kÿ** nek-ÿ **ne** engheg-gy-en
just who DAT-3SG not yield-SBJV-3SG
'It (the devil's temptation) secures greater merit for the good-natured man; just **nobody** should yield to it'
(OH, Érdy Codex, 1524: 82b)

d. En-nek feletth-e ÿnden-ek kÿk hall-ÿak wal-a **kÿ** **mÿnd**
this-DAT above-POSS.3SG everyone-PL who-PL hear-3PL be-PST who each
ew nÿelw-ek-en oketh zol-wan ...
they tongue-POSS.3PL-SUP them speak-CVB
'Furthermore, everyone who heard them, **each** (hearing them) speaking in his own tongue ...'
(OH, Érsekújvári Codex, 1529: 80v)

Hungarian indeterminate pronouns gradually developed specific prefixes in their non-interrogative functions, as illustrated in (18).

(18)

interrogative:	*ki*	'who'
relative:	*a-ki*	'who'
indefinite:	*vala-ki*	'somebody'
negative:	*sen-ki*	'nobody'
universal:	*minden-ki*	'everybody'
free choice:	*bár-ki*	'anybody'

The formal specification of indeterminate pronouns functioning as relatives started in the Late Old Hungarian period, but it was a long, gradual process, hence a bare pronoun in the initial clause of Old Hungarian correlative constructions could, in principle, be either an indeterminate pronoun or a relative in Spec,CP. However, the bare pronoun regularly alternates with its indefinite counterpart, which cannot be categorized as a relative pronoun. Compare:

(19) a. **ký** nem wew-tth-e hewsa[a]g-ba eew lelke-th, **eez** weez-en aldomas-th
who not take-PST-3SG vanity-to his soul-ACC this take-s pledge-ACC
wr-thol
Lord-from
'**Who** didn't take his soul to vanity, **he** takes pledge from the Lord.'
(OH, Festetics Codex, 1494: 13)

b. **vala-ki** akar-and len-nÿ ez velag-nak barat-t'a, [**az**] isten-nek
some-body want-FUT.3SG be-INF this world-DAT friend-3SG that god-DAT
lez-en ellenseg-e.
be-3SG enemy-3SG
'**Who** wants to be the friend of this world, **he** will be the enemy of God.'
(OH, Bod Codex, 1500–1525: 2v)

(20)a. **my-th** eghzer meg_ zerz-ett-el ... [**az-t**] tewbzer nem kel hoz-yad
what-ACC once obtain-PST-2SG that-ACC more not need ALLAT-2SG
we-nn-ed
take-INF-2SG
'**What** you have once obtained ... **that** you don't need to take again for you.'
(OH, Jogi szabály, 1476)

b. **vala-mi-t** èn-nèk-ėm mond-ād-az [**az-t**] tėz-em te nèk-ėd
some-thing-ACC I-DAT-1SG say-FUT-2SG that-ACC do.1SG you DAT-2SG
'**What** you tell me, **that** I will do for you.'
(OH, Bécsi Codex, 1416: 1.7)

(21)a. Az leelek **hol** akar, **oth** lehel.
the soul where want.3SG there breathe.3SG
'**Where** the soul wants to breathe, **there** it breathes.'
(OH, Jordánszky Codex, 1516: 630)

b. **vala-hol** kètt-èn ag̃ harm-an egbe gǫlèkez-ènd-n^{c}
some-where two-ADV or three-ADV together assemble-FUT-3PL
èn nèu-ē-bè, **ot** èn ǫ-kǫzǫtt-ǫc vagoc
I name-1SG-in there I they-among-3PL am
'**Where** two or three persons gather in my name, **there** I am present among them.'
(OH, Müncheni Codex, 1416/1466: 24v)

We also find the indefinite version of the indeterminate pronoun/proadverb in sentences where the demonstrative of the second clause is implicit, which supports the assimilation of this clause type to the pattern in (19)–(21):

(22) **Vala-houa** men-ènd-ėz [oda] mėg-ėc 7
some-where.to go-FUT-2SG [there.to] go-1SG &
vala-hol lakoz-ād-al èn es [ott] lakoz-om
some-where live-FUT-2SG I also [there] live-1SG
'**where** you go, [there] I will go, and **where** you live, [there] I will live'
(OH, Bécsi Codex 1416: 1.2)

Though Old Hungarian constructions with an indeterminate or indefinite pronoun in their initial clause are formally similar to the Ob-Ugric constructions discussed in Section 3.1, they are interpreted differently; they have a free choice reading with a universal force. Free choice pronouns are analyzed as indefinites that assume a universal force in generic and modal contexts (Giannakidou 2001; Kratzer and Shimoyama 2002; Halm 2016). This is what we see in (19)–(22). The generic/modal context is occasionally made explicit in the form of a conditional complementizer (23a,b). The universal reading arising in generic and modal contexts is occasionally made explicit by a universal determiner (24a,b).

(23) a. **ha ki** kèdeg le-tèrièzked-uen nem imad-ang-a, [**az**] azon v̇dọ-bèn
if who in.turn down-lie-CVB not adore-FUT-OBJ.3SG that that time-in
èrèz-tèt-ic az egọ tùz-nèc kèmencè-ie-bè
let-PASS-3SG the burning fire-DAT furnace-POSS-into
'**If somebody** lying down will not adore him, **he** will be let into the furnace of burning fire.'
(OH, Bécsi Codex, 1416: 2.125)

b. **Ha ki** èi-ènd-ic è kèńer-bọl [**az**] ėl ọrọcke
if who eat-FUT-3SG this bread-from that lives forever
'**If somebody** eats from this bread, **he** lives forever.'
(OH, Müncheni Codex, 1416/1466: 91r)

(24) a. **Mynden vala ky** Vr-nak new-ee-th hyw-angy-a, **az** ydwọ́zẅl.
every some who Lord-DAT name-POSS-ACC call-FUT-3SG that is.saved
'**Whosoever** shall call upon the name of the Lord, **he** shall be saved.'
(OH, Jordánszky Codex, 1516: 711)

b. **Menden vala ki** vall-and èngemèt èmberec èlọt. èmber-n^{c} fi-a
all some who confess-FUT.3SG me people before man-DAT son-POSS
val'l'a **ọtèt** istèn-n^{c} angal-i èlọt.
confesses him God-DAT angel-POSS.PL before
'**Whosoever** shall confess me before men, **him** shall the Son of man also confess before the angels of God.'
(OH, Müncheni Codex, 1416: 70r)

In (23) and (24) the indeterminate/indefinite pronoun is not in Spec,CP yet. In (23), it is preceded by the conditional complementizer in C. As for (24a,b), universal quantifiers undergo overt Q-raising in Hungarian; they land in post-topic position, left-adjoined to the predicate phrase (TP or FocP). It is unclear whether the Q-raised universal is *menden*, or *menden vala ki*; in either case, the *wh*-pronoun is below/after the CP domain. This fact indicates that the Old Hungarian correlative pattern – at least the (b) examples in (19)–(21), and examples (22)–(24) – are not canonical correlative constructions yet. At the same time, they are closer to the canonical correlative pattern than their Ob-Ugric counterparts. Whereas the likely reading of the indeterminate pronouns in the Khanty and Mansi examples in (8)–(13) and (16) is an existentially bound indefinite reading, the Hungarian indeterminate and indefinite pronouns in (19)–(22) have a free choice interpretation with a universal force. Whereas the two clauses of the Khanty and Mansi constructions are in an asyndetic or syndetic coordinate relation, the relation of the two clauses of the above Old Hungarian constructions is not symmetric any more; the initial clause is deranked as compared to the second clause, hence no conjunctive particle can appear between them. These differences suggest that the Ob-Ugric and Old Hungarian correlative patterns represent two subsequent stages of the grammaticalization of the correlative construction.

Actually, Hungarian has preserved a relic on the basis of which the transition between the two stages can be presumed. Two determiners, *vala-hány* 'some-how.many', and *vala-mennyi* 'some-how.much' can function both as assertive existential indefinites and as indefinites bound by a covert generic/modal operator in Modern Hungarian, as well, hence we can replicate with them how the coordinate structure corresponding to stage 1 (25a) could be

reanalyzed as a subordinate structure corresponding to stage 2 (25b). (25c) is the canonical correlative version, with a relative pronoun in Spec,CP.

(25)a. Existentially bound indefinite pronoun, parataxis or coordination:
(**Vala-hány könyv-et**) pro$_i$ megve-tt (**vala-hány könyv-et**), (és)
some-many book-ACC buy-PST.3SG (some-many book-ACC) and
az-ok-at Mari$_j$ mind elolvas-t-a.
that-PL-ACC Mary all read-PST-OBJ.3SG
'She$_i$ bought **some books**, (and) Mary$_j$ read them all.'

b. Generically bound indefinite pronoun, subordination:
Vala-hány könyv-et pro$_i$ megve-tt, **az(-ok)-at** Mari$_{i/j}$ mind elolvas-t-a.
some-many book-ACC buy-PST.3SG that-PL-ACC Mary all read-PST-OBJ.3SG
'**Whatever number** of books she$_i$ bought, Mary$_{i/j}$ read **them** all.'

c. Relative pronoun, subordination:
A-hány könyv-et pro$_i$ megve-tt, **az(-ok)-at** Mari$_{i/j}$ mind elolvas-t-a.
some-many book-ACC buy-PST.3SG that-PL-ACC Mary all read-PST-OBJ.3SG
'**Whatever number** of books she$_i$ bought, Mary$_{i/j}$ read **them** all.'

In the initial construction in (25a), the indefinite phrase – in situ or preposed into topic position – is bound by default existential closure. The two clauses are in a symmetric relation, and a coordinating conjunction can also appear between them. As is predicted for coordinate clauses (Langacker 1969), the pronominal subject of the first sentence cannot be coreferent with the lexical subject of the second sentence. In (25b), the indefinite phrase is bound by a covert generic quantifier. As expected in subordination structures, the pronominal subject of the subordinate clause can be coreferent with the lexical subject of the main clause. (25c) already contains a relative pronoun in Spec,CP. The asyndetic/paratactic version of (25a), the pattern also attested in Ob-Ugric, is formally identical with the correlative construction in (25b). (Actually, whereas the demonstrative is in the plural in (25a), it can also be in the singular in (25b,c); in fact, singular is the preferred, default option. This follows if the demonstrative is coreferent with a group-denoting expression in (25a) and with a bound variable in (25b).) In Old Hungarian, a similar formal identity must have existed between the existentially bound version and the version bound by a generic/modal operator in the case of all indeterminate pronouns. It is this formal identity which must have made the transition from (25a) to (25b) possible.[2]

The constructions in (25a–c) outline a developmental path from parataxis to the correlative construction. Interestingly, Belyaev and Haug (2020) argue for a similar evolution based on theoretical considerations, supported by some Early Latin and Hittite data.

3.3 Belyaev and Haug's (2020) theory of the evolution of correlatives

Belyaev and Haug's study aims to answer the question what the source of the universal/free choice interpretation of the relative pronoun of the correlative clause is. The proposed answer is admittedly speculative; nevertheless, it shows surprising correspondence with the grammaticalization path attested in the Ugric languages.

[2] The question may arise why *vala-hány* 'some-how.many' and *vala-mennyi* 'some-how.much' of all indefinite pronouns have preserved their ability to function as variables bound by a generic operator. Presumably because *hány* and *mennyi* are the least frequent of all *wh*-pronouns, therefore, they were the last to develop specific relative forms. Whereas the first known occurrence of the relative *a-ki* 'who' is from 1372, the first occurrence of the relative *a-mennyi* 'how much' is from 1509, and the first occurrence of *a-hány* 'how many' is from 1819.

Belyaev and Haug observe that the wh-pronouns of correlative clauses function in many languages as interrogatives and indefinites, and hypothesize that originally they were indefinites in paratactic constructions. They note that this has already been raised in connection with two early Indo-European languages, Latin and Hittite. They cite Kroll (1910), who argued about Early Latin correlative constructions that in many cases they are paratactically linked sentences, and what seems to be a relative pronoun in them is, in fact, an indefinite. Observe the German translation that Kroll assigned to (25):

(25) **quae pequnia** recepta erit, **ea pequnia** emere conducere locare dare...
some money received will_be, that money to_buy to_hire to_rent
liceto
can_be_used
'irgend welches Geld will einkommen; mit diesem Gelde soll es erlaubt sein zu kaufen...'
[Some money will be received; this money can be used to buy, hire and rent]
(Latin, Corpus Inscriptionum Latinarum 1: 603, Lex Furfensis 58, cited by Kroll 1910: 8)

Belyaev and Haug also refer to Hahn (1946) in passing. Hahn (1946) argued that in the first Hittite documents the relative pronoun of correlative clauses is indistinguishable from an indefinite pronoun as regards its form and position, and is ambiguous as regards its function. Furthermore, in 90% of the 390 apparent correlative constructions analyzed by her, the conjunctive particle *nu* appears between the clauses.[3] Hahn concludes that most of the 390 sentences have two equally likely interpretations: a coordinate and a correlative reading – see e.g. (26a), glossed by Lehmann (1980). In several examples it is the coordinate reading with an indefinite pronoun in the initial clause that yields a more coherent interpretation. In (26b), for example, there is a number mismatch between the pronouns of the two clauses, which would be atypical of a correlative construction.

(26) a. GUD-*pūḫugarin-ma* ***kuedani*** **UD-*ti*** *únuēr* ***nu-za*** dUTU-*ši*
ox-scape PTC which-on day-on decorated.3PL PTC PTC sungod
apēdani **UD-*ti*** *warapta*
that-on day-on he bathed
'**On which day** they adorned the scapeox, **on the day** the sungod bathed.'
(Hittite, Lehmann 1980: 133)

b. *nam-ma-ma* ***ku-i-e-es*** LÚ.MEŠŠANGA *nu-za* ***ku-is*** *ŠA* KÁ *e-es-du*[4]
?'(those) **who** are priests let **one** be doorkeeper'
or: '**some** are priests, let **one** be doorkeeper'
(Hittite, Hahn 1946: 77)

Recently, Motter (2023) has provided syntactic evidence demonstrating that the clauses of the Hittite correlative construction are in a paratactic relation, with *nu* acting as a discourse connective.

Belyaev and Haug (2020) analyze the *wh*-pronouns of correlative clauses as Hamblin-indefinites that are bound by a higher (generic or modal) operator. A condition of the

[3] According to Hoffner and Melchert's Hittite grammar (2008: 390-391), the particle *nu* in Old Hittite could connect two independent clauses, two subordinate clauses, as well as a *when*-clause and a following independent clause, and a preposed relative clause and a main clause – i.e., the seemingly subordinate contexts where it occurred are those where the clause containing the wh-pronoun may originally have been a paratactic sentence with an indefinite pronoun.

[4] Hahn (1946) does not gloss her examples.

emergence of a quantificational interpretation in the antecedent of the correlative construction is that there be a conditional semantic relation between the two paratactically linked sentences (which is a crosslinguistically common phenomenon – cf. *You drink one more beer and I leave*). A further ingredient is an anaphoric connection between the clauses, which allows the assumption of a topic-comment relation between the two sentences, with the demonstrative in the second clause functioning as a topic doubling element. It is the topic-comment structure that allows the assumption of a covert generic or modal operator in a tripartite structure (Partee 1991), with the topic clause functioning as the restrictor and the comment clause interpreted as the nuclear scope (see Keshet 2013: 225-232).

The Ugric languages provide empirical evidence for the hypothetical scenario outlined by Belyaev and Haug. Ob-Ugric examples of the type illustrated in (3)–(9), involving two independent sentences linked paratactically or by means of a coordinating conjunction, displaying an anaphoric relation, represent the initial stage of the evolution. The indeterminate pronoun in the initial clause is an existentially bound indefinite with no universal force. Many Old Hungarian examples still contain an indefinite pronoun in the first clause of the construction; however, the pronoun already has a free choice or universal interpretation, i.e., it is clearly bound by an overt or covert conditional or generic operator. This is the interpretation of the correlative construction in Modern Hungarian, as well, but the pronoun of the initial clause is already a relative pronoun that occupies Spec,CP.

Hahn's (1946) Hittite data and Kroll's (2010) Early Latin data in most cases show the same ambiguity that is illustrated by the Hungarian (25a,b), i.e., they can be interpreted either as coordinate sentences with an indefinite pronoun in the initial clause, or as correlative constructions involving a covert quantifier. The reanalysis of the former paratactic pattern as a quantificational structure must have been the key momentum of the emergence of the correlative construction not only in the Ugric languages but also in Early Indo-European.

4. From correlatives to adnominal relative clauses

The data surveyed in Section 3 suggest that the first type of finite relative clause to emerge in the Ugric languages was the correlative clause. In Old Hungarian, we can also follow its further evolution; it was the source of the development of adnominal finite relative clauses.

This process seems to have been triggered as part of a general drift from head-final to head-initial syntax. The directionality change also reversed the default order of the correlative clause and main clause – as illustrated e.g. by example (27), where the correlative status of the subordinate clause is shown by the presence of two relative pronouns in it:

(27) **ki-ki** mind [**ott**] le yl-e, [**akinek** **hol** helie vala]
who-who all [there] down sit-PST.3SG who-DAT where place-POSS was
'**Whereever each one** found a place, **he** sat down [**there**].'
(OH, Lobkowicz Codex, 1514: 178)

In the reverse order version, the bare indeterminate pronoun is often right-adjacent to its demonstrative correlate, as is the case in (28). This must have been the structure in which the subordinate clause was reanalyzed as the modifier of the demonstrative, i.e., as a so-called light-headed relative clause. Notice that the frequent *vala-* prefix of the correlative pronoun is practically never present in such constructions.

(28) Byzonyaual bodog-og **ezek** [**kyk** meg ezen feld-en hasonlatos-ok az
surely blessed-PL these who-PL still this earth-on similar-PL the
angel-ok-hoz]
angel-PL-ALLAT

‘**These who** are similar to the angels still on this earth are surely blessed.’
(OH, Lobkowicz Codex, 1514: 191)

The fact that the relative clause belongs to the main clause demonstrative is especially obvious when the demonstrative + relative clause string is both preceded and followed by main clause material. In (29), the constituency of the demonstrative + relative clause strings is also shown by their being coordinated.

(29) embery bewlczeíeg-nek zem-e-y nem-czak [**azok-ot** [**ky-k** belewl]]
human wisdom-DAT eye-POSS-PL not-only those-ACC who-PL inside
de [**azok-ot** [**ky-k** kyuel vad-nak]] es meg mutatt-yak
but those-ACC who-PL outside be-3PL too PRT show-3PL
‘The eyes of human wisdom show not only **those who** are inside but also those who are outside.’
(OH, Jókai Codex, 1372: 129)

In many cases there are no formal criteria to decide whether the subordinate clause is a postposed correlative or a relative clause already. Thus the subordinate clause in (30a) could be either a postposed correlative or an extraposed light-headed relative (30a). The subordinate clause in (30b) could be either a postposed correlative with a pro-dropped main clause correlate, or an adnominal or headless relative. In such cases, the criterion distinguishing correlatives from adnominal and headless relatives is semantic: owing to their covert generic or modal operator, correlatives have a universal force. (Based on this criterion, (30a) involves a correlative clause, and (30b), an adnominal or free restrictive relative clause.)

(30) a. Es ne akar-y-atok **azok-at** fel-ny, [**ky-k** az teſt-et megh ǫ́l-yk].
and not want-IMP-2PL those-ACC fear-INF who-PL the body-ACC PRT kill-3PL
‘And don't want to fear **those who** kill the body.’
(OH, Jordánszky Codex, 1514: 384)

b. keer-gy-ed azok-at ky-k hall-ott-ak [**pro** [**My-t** en zol-t-am]].
ask-IMP-2SG those-ACC who-PL hear-PST.3PL what-ACC I say-PST.1SG
‘Ask those who heard **what** I said.’
(OH, Apor Codex, 1416: 212)

The shift from the correlative to the relative construction is also difficult to point out in the case of locative and temporal relative clauses, as the default position of locative and temporal adjuncts is in the left periphery of the Hungarian sentence; furthermore, it seems incidental whether an explicit locative or temporal correlate appears in the main clause. We can again rely on the semantic criterion, according to which (31a) involves a correlative clause, and (31b), a left-dislocated relative clause:

(31) a. [**hol** az zerelm] **ot** az zem
where the love there the eye
‘**Where** there is love, **there** is eye (to see).’
(OH, Cornides Codex, 1514: 55r)

b. [**hol** yl-ual-a]$_i$ [**ottan** t_i] germekek gylekez-e-nek hozzaya sok-an
where sit.3SG-be-PST there children gather-PST-3PL to.him many-ADV
‘**Where** he was sitting, there children were gathering around him in great numbers.’

(OH, Jókai Codex, 1372: 21)

Temporal correlatives appear to be non-distinct from conditional clauses, the complementizer of which, *ha* ‘if’, grammaticalized from an Old Hungarian indeterminate temporal adverb (cf. *vala-ha* ‘sometime’, *so-ha* ‘at no time’). *Ha* often co-occurs with the main clause correlate *tahat* ‘then’(32a). Temporal relative clauses introduced by the pronoun *mi* ‘what’ supplied with a temporal case or postposition mostly have episodic readings (32b,c).

(32) a. **Ha** az soror-ok kevzzvl valamelyk-nek at-ya an-ya ... meg
if the soror-PL from.among some-DAT father-POSS mother-POSS PRT

hal val-a **Tahat** ez zent zvz vele evzve sirat-ya val-a
die.3SG be-PST then this saint virgin with.her together mourn-3SG be-PST

‘If the father or mother of any of the sorors would die, then this saint virgin would mourn together with her’
(OH, Margit legenda, 1510: 36)

b. [**Mi-koron** hat-tal fordul-t vol-na hozzaÿa] lat-a
what-time.at back-with turn-PERF.3SG be-COND to.him see-PST.3SG
az vitez o benn-e kilemb kilemb kin-ok-nak nem-i-t
the knight he IN-3SG different torment-PL-DAT sort-POSSPL-ACC
‘**When** he turned to him with his back, the knight saw sorts of different torments in him.’
(OH, Bod Codex, 1500–1525: 3v)

c. [**Mi vtan** kedig mėgfordol-t-uol-na o-nap-a-hoz]
what after however turn-PERF.3SG-be-COND she-mother.in.law-3SG-to
[**az után**] hall-a tǫll-ǫ ...
[that after] hear-PST.3SG from-2sg
‘**Whereafter** she had turned to her mother in law, **thereafter** she heard from her ...’
(OH, Bécsi Codex, 1416: 6)

In constructions like (29), involving a demonstrative modified by a relative clause, the main clause distal demonstrative came to be reanalyzed as part of the relative pronoun (see Bacskai-Atkari and Dékány 2014; 2020). Whereas a main clause demonstrative modified by a relative clause shares the singular or plural number of the relative phrase, and bears the case assigned to it in the main clause, a demonstrative merged with the relative pronoun is not marked either for number or for case (33a). Eventually it also lost its final *z* sound. The incorporation of the demonstrative into the relative pronoun was followed by the renewal of the main clause demonstrative (33b).

(33) a. *Rebracketing*:
laang meg egeth-e [**az_kÿk** benes-ek val-a-nak]]
flame PRT burn-PST.OBJ.3SG who sinful- PL be-PST-3PL
‘The flame burned them who were sinful.’
(Middle H, Kulcsár Codex, 1539: 131r)

b. *Erosion of original* az*; demonstrative renewal*:
Ez [**az** [**a'ky** felő̊l meg vagyon irua]]

this that who about PRT is written
'This is [that [of whom it is written]]'
(Middle H, Károli, 1590: 57v)

The main clause demonstrative originating in the correlative construction not only survived as the head of light-headed relative clauses; it is also present in lexically headed relatives. Its use is a relic of the correlative construction in which the relativized NP is spelled out in both correlates, e.g.:

(34) **melÿ** **tag-ok-nak** **mÿatta** vetkez-et val-a **azon** **tag-og-ba**
which member-PL-DAT because.of sin-PERF.3SG be-PST those member-PL-in
gevtr-et-ÿk val-a es
torment-PASS-3SG be-PST PRT
'Which members she had sinned for, she was being tormented in those members.'
(OH, Példák könyve, 1510: 40)

When the order of the relative clause and the main clause was reversed, the lexical element of the relative phrase could be spelled out after the *wh*-determiner (35a), or it could be ellipted (35b), but the most common solution was the replacement of the whole noun phrase with a relative pronoun (35c):

(35) a. lee-n eredethy [**azon ÿdó-ben** ... [**mÿ** **ÿdóben** esÿk Zent Marghÿt
be-PST.3SG origin-3SG that time-in what time-in falls Saint Margaret
azzon nap-ÿa]]
lady day-POSS
'Its origin was **on that day on which day** Lady Saint Margaret's day is.'
(OH, Érdy Codex, 1524: 239)

b. ha mėǵmarad-and-atoc [**ė** **mv̇uèlkėdet-ėc-bèn** [**mėll-Ø-èt** kèzd-ėtt-ètec]]
if remain-FUT-2PL this action-PL-in which-ACC start-PST-2PL
'If you remain **in the activities which** you started'
(OH, Bécsi Codex, 1416: 16)

c. ne rəstel-i-etek [**azon za-bol** oruossagot mōd-ani [**ki-bøl** seb-ək
not loth-IMP-2PL that mouth-from cure-ACC say-INF what-from wound-PL
lo-tt-enek]][5]
become-PST-3PL
'Don't loth to say cure from **that mouth that** wounds came from.'
(OH, Birk Codex, 1474: 2)

These are the three patterns of the restrictive relativization of lexical elements in Modern Hungarian, as well, e.g.:

(36) a. az a ház, **amely** **ház-ban** fel-nő-tt-em
that the house which house-in up-grow-PST-1SG
'the house in which house I grew up'
b. az a ház, **amely-ben** felnőttem
'the house in which I grew up'

[5] The pronoun *ki* was [+/-human] in Old Hungarian; in Modern Hungarian it is [+human].

c. az a ház, **ahol** felnőttem
‘the house where I grew up’

Adnominal relative clauses have also appeared in Khanty. As shown by Dékány et al. (2020), finite relative clauses introduced by a *wh*-based relative pronoun have emerged in all major Khanty dialects (on the Northern Dialect, see Nikolaeva 1999: 45; on the Eastern Surgut dialect, see Csepregi 2012; on the easternmost Vasyugan dialect, see Filchenko 2010: 497). Adnominal finite relative clauses are always postnominal, e.g.:

(37) **Puγəł**, [**mətapi-nə** ma säm-a pit-əm], ənəł łår qånəŋ-nə åməs-ł.
village which-LOC I eye-LAT fall-PST.3SG big lake shore-LOC sit-PRS.3SG
‘**The village in which** I was born (lit. fell into eye) is located on the shore of a big lake.’
(Eastern Khanty, Csepregi 2012: 88)

Dékány et al. identify the correlative construction as the starting point of the grammaticalization of Khanty finite subordination. They argue for the following developmental path:

(38) correlative > free relative/light-headed relative > lexically headed finite relative

The cut-off point after correlatives is motivated diachronically: in traditional Khanty texts we only find constructions interpretable as correlatives but there are no free relatives or light-headed relatives. The cut-off point before lexically headed relatives is motivated by the fact that the set of *wh*-pronouns occurring in free relatives and light-headed relatives is larger than the set of relatives occurring in lexically headed finite relatives; *qŏjaγi* ‘who’ and *müwəłi* ‘what’ have not made it to the last stage of the grammaticalization path, as shown by the minimal pair in (39):

(39)a. **Əj məta qo** pə, [**qŏjaγi** Surgut-nə wăł-ł], mant əntə wu-ł.
one some man PRT who Surgut-LOC live-PRS.3SG I.ACC not know-PRS.3SG
‘**No one who** lives in Surgut knows me.’
(Khanty, Dékány et al. 2020: 36)

b. **Ŏnəltətə ne,** [**mətapi/*qŏjaγi** Surgut-nə qăntək köł ŏnəltə-ł],
teaching woman which/who Surgut-LOC Khanty language teach-PRS.3SG
tem qătəł-nə ńyńť-əł-0.
this day-LOC rest-PRS.3SG
‘**The woman teacher which/*who** teaches Khanty in Surgut is resting today.’
(Khanty, Dékány et al. 2020: 28)

In Hungarian the set of relative pronouns found in free relative/light-headed relative clauses is identical with the set found in lexically headed relatives, hence there is no indication that emergence of light-headed relative clauses preceded the emergence of lexically headed the relatives.

Dékány et al. (2020) regard the emergence of finite adnominal relative clauses in Khanty as contact-induced grammaticalization triggered by Russian influence. They describe the change as "wholesale borrowing of the Russian structure", which includes: (i) order with respect to the head noun, (ii) finite verb inflection, (iii) syncretism between interrogative and relative pronouns, (iv) high preference for the relative pronouns to be the first element in the relative clause, and (v) the exclusion of the equivalents of *who* and *what* from lexically

headed finite relative clauses but not from light-headed finite relatives, free relatives and correlatives. As Khanty replicated the interrogative-relative syncretism of Russian without borrowing lexical material, and as relative pronouns first appeared in correlatives and were extended to externally headed RCs only later, replica grammaticalization happened.

Actually, if the starting point of this grammaticalization path was the paratactic construction with an indeterminate pronoun, the correlative construction could also evolve without replicating an external pattern, as is predicted by Belyaev and Haug (2017), and as presumably happened in Hungarian. In that case, the further evolution of finite relative clauses could also proceed as a basically endogenous process, similar to that observed in Hungarian. The corresponding Russian constructions could play a role in confirming and accelerating the process.

Similar evolutionary paths have been argued for in early Indo-European, as well. In Hittite, the source construction was the same correlative construction as in Ugric, with an indefinite pronoun grammaticalized into a relativizer (Justus 1976; Lehmann 1980: 130–135). The recategorization of the indefinite as a relative pronoun has been related to a shift in the order of the relative clause and the main clause, and, more generally to a shift from OV to VO (Lehmann 1980: 138–139). The Hittite relativized phrase, too, could be either repeated, or pronominalized, or or pro-dropped in the main clause, and these options could develop into lexically headed, light-headed and headless relatives respectively. Ram-Prasad (2022: 172) argues based on Hittite, Greak and Sanskrit examples that the reanalysis of postposed correlatives as adjoining to or being a complement of a main clause head noun, rather than adjoining to the matrix CP, was nascent in Proto-Indo-European already.

5. From the correlative construction to the complement clause

Hungarian, unlike the Ob-Ugric languages, has developed not only finite relative clauses but finite complement clauses, as well. The Hungarian general complementizer *hogy* ‘that’ has been traced back to the the relative pronoun *hogy* ‘how’ of correlative construction (É. Kiss and Gugán 2021; É. Kiss 2023b).

Hogy is claimed to have grammaticalized into a complementizer in the context of reporting verbs selecting a propositional complement. In SOV Proto-Hungarian, and frequently still in early Old Hungarian, too, the propositional complements of *say*-verbs, having independent tense, mood, and force, were formulated paratactically, as independent sentences – see (40), and for more examples and for a more detailed discussion, É. Kiss (2023b).

(40) En kedeg **vő mond-ok** ty-nek-tek: Zeres-s-etek ty ellensegh-tek-et
I in_turn so say-1SG you-DAT-2PL love-IMP-2PL you enemy-2PL-ACC
‘I, in turn, **say so** to you: Love your enemy’
(OH, Jordánszky Codex, 1514: 368)

As in example (40), the reporting verb was often accompanied by the cataphoric demonstrative *úgy* ‘so’, announcing the propositional complement to be reported in the following sentence. (As argued by Munro (1982), *say*-verbs are weakly transitive crosslinguistically; hence their complement is not necessarily realized as an object; it is often referred to by a proadverb of the type *úgy* ‘so’). The string *úgy mond* ‘says so’ was so common in Old Hungarian that it grammaticalized into the discourse particle *úgymond* ‘so to say’ (Dömötör 2008).

The drift from head-final to head-initial syntax in Proto-Hungarian and Early Old Hungarian brought along a drift from non-finite embedded clauses (with a clause-final participial, infinitival or nominalizing suffix acting as the subordinator) to finite embedded clauses with an initial complementizer (É. Kiss 2023a). When pressure emerged to integrate

the propositional complements of *say*-verbs as finite subordinate clauses, the only subordinators that Hungarian had were the relative pronouns of the correlative construction. *Úgy* 'so', the demonstrative companion of *say*-verbs in paratactic constructions, evoked its correlate *hogy* 'how' from among the set of relative pronouns. In manner correlative constructions, *úgy* was the pair of the relatives *hogy* 'how' and *mi-ként* 'how' lit. 'what-ESSIVE', as illustrated by a canonical correlative construction in (41a) and a reverse order version in (41b):

(41) a. mert te vrā [**mikēt** akar-t-ad] **ug** tọ-tt-èl
because you lord.1SG how want-PST-OBJ.2SG so do-PST-2SG
'because you my lord, **how** you wanted **so** you did'
(OH, Bécsi Codex, 1416: 240)

b. furiscte mus-i etet-ý ýmlet-i **ug** [**hug** ana
bathe-OBJ.3SG wash-OBJ.3SG feed-OBJ.3SG nurse-OBJ.3SG so how mother
scilutt-e-t]
offspring-POSS-ACC
'She bathes, washes, feeds, nurses him **so how** a mother her offspring.'
(OH, Königsbergi töredék, 14th c.)

Apparently, the simple *hogy* was more apt to grammaticalize into a complementizer than *miként*, bearing the suffix of a semantic case. The *úgy...hogy* 'so...how', 'so...as' pair was also more common as it was used in consecutive constructions:

(42) **vgy** uala ewn-belewle kywl **hogy** te bezzed-y-t-hez... semmy-t
so was himself-from beside **as** you speech-POSSPL-2SG-to nothing-ACC
nem hayl-a
not incline-PST.3SG
'He was **so** (much) beside himself **as/that** he was not inclined to your speeches at all'
(OH, Jókai Codex 1372: 10)

Consequently, when the drift to SVO syntax brought along a shift to head-initial subordination, and the language adopted a subordinator from the correlative constructions, the cataphoric *úgy* elicited the use of *hogy*:

(43) Mert **vgy** **mond-ok** ty-nek-tek, **hogy** hatalm-a vaǧon Iſten-nek
for so say-1SG you-DAT-2PL that might-3SG is God-DAT
ez kọ́wek-bọ́l fel tamazt-any Abraham-nak fy-a-y-t
this -stones-from up resurrect-INF Abraham-DAT son-POSS-PL-ACC
'For I **say so** to you **that** God has power to resurrect Abraham's sons from these stones.'
(OH, Jordánszky Codex, 1516: 361)

In (41a,b), the relative pronoun is a phrasal category with lexical features (the synonym of *milyen módon* 'in what manner') preposed to Spec,CP from a VP-adjoined position. In (43), by contrast, it has already lost its modal content and its phrasal structure; it is a complementizer base-generated in C. Nevertheless, the *hogy*-clause still preserves the adjoined status of correlative clauses; it does not elicit agreement on the matrix verb. This construction is akin to the so-called explicative structure (Hettrich 1988: 395), in which the complement clause explicates the content of its pronominal (in this case, proadverbial) correlate.

The integration of the complement clause into the matrix VP becomes complete when an object agreement morpheme appears on the matrix verb:

(44) Bijzon **mond-om** nek-ed, [hogij ma, ez eijel mij-nek elewtt-e
truly say-**OBJ.1SG** DAT-2SG that today this night what-DAT before-POSS
kakas kerczer enekel, harōzor taga-cz meg engemet]
cock twice sing-3SG thrice deny-2SG PRT me
'Verily, I say to you that today, this night, before the cock crows twice, you shall deny me thrice'
(MH, Pesti 1536: 59r)

When *hogy*-clauses spread from the context of reporting verbs to other contexts with a propositional complement, among them the contexts of predicates assigning oblique cases to their complements, a demonstrative appeared in the main clause to pick up the case assigned by the matrix verb (45a). The demonstrative also came to be used as a place holder for the subordinate clause in positions not accessible to clausal constituent, e.g., in the preverbal focus slot (45b).

(45) a. biz-ván **ab-bā**, hogy a' ki kezd-ett-e bennetek a' jó dolg-ot,
trust-CVB that-in that who start-PST-3SG INES-2PL the good thing-ACC
el-végezi a' Kriſtus Jéſus nap-já-ig
completes the Christ Jesus day-POSS-until
'believing in it that he who has begun a good work in you will perform it until the day of Jesus Christ'
(MH, Káldi 1626: 1084)

b. deh **ćzak az** le-gÿ-en ọrọm-ẃnk: [hoǵh ffel ÿr-attat-oth
but only that be-SJV-3SG joy-1PL that up write-PASS-PST.3SG
meń orzag-ban az mÿ new-ẃnk]
heaven-in the we name-1PL
'but **only that** shall be our joy that our name has been written in heaven'
(OH, Könyvecse, 1521: 2r)

The grammaticalization path of *hogy* from a relative proadverb of manner correlative clauses to the general complementizer of complement clauses can be summed up as follows: *Hogy,* originally meaning 'how', evolved into a complementizer meaning 'that' in the context of *say*-verbs. In Proto-Hungarian and often in Early Old Hungarian, too, the propositional complements of *say*-verbs were paratactic sentences announced in the reporting sentence by the cataphoric proadverb *úgy* 'so'. Paratactic complements came to be integrated into the reporting sentence by means of a subordinator borrowed from correlative constructions, and *úgy* evoked the use of its correlate *hogy* from among the set of *wh*-relatives. *Hogy* employed as a subordinator lost its lexical features; it became a head base-generated in C. Complement clauses introduced by *hogy* were first adjoined to the main clause, and later became integrated into the matrix VP. Eventually *hogy*-complement clauses developed pronominal main clause associates to pick up the case assigned to them in the main clause and/or to represent them in main clause operator positions.

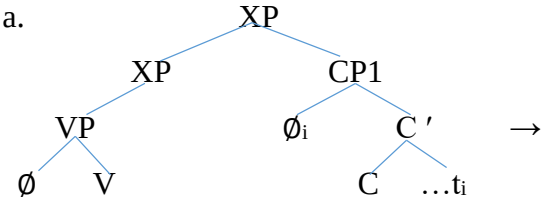
We find complementizers cognate with relative pronouns in Indo-European, as well, among them in the Romance, Slavic, and Germanic languages, but the details of their grammaticalization are less clear. Axel-Tober (2017) has outlined a grammaticalization path for the Germanic *that*-type complementizers. She claims that the German declarative

complementizer *dass* developed from a correlative construction, but she only traces it back to the explicative clause illustrated in (46):

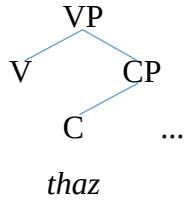
(46) Sliumo ságeta er mo tház, \ **tház** er mo er kúnd was,
suddenly said he him that that he him before known was
'He (= Jesus) told him at once that he knew him already'
(Old High German, Otfrid II 7.61, cited by Axel-Tober 2017: e40)

Adopting Haegeman's (2012: 273) analysis, Axel-Tober regards the subordinate clause of (46) as a (cor-)relative clause, which contains an empty relative operator in Spec,CP that relativizes the implicit event argument of the main clause. *Thaz* is assumed to be the C head of the CP harboring the relative operator. The reanalysis of *thaz* as a declarative complementizer was made possible by the disappearance of the relative operator from Spec,CP, which must have been a consequence of the fact that its main clause correlate was often a silent pronominal that could be ignored. The hypothesized structural change is represented in (47):

(47) a.



→ b.



Structure (47a), however, is problematic in at least two respects. First, it is unclear how *thaz*, an uninflected C head, relates to the inflected *thaz* functioning as a relative pronoun. If it was originally an inflected relative in Spec,CP, it is unclear why it was forced out of Spec,CP by an invisible relative operator, or if it was lowered from Spec,CP to C, why Spec,CP had to be filled again. Second, a hallmark of the correlative construction is an anaphoric relation between the relative operator and its main clause correlate, but in (46) we find no such relation between the operator relativizing the event argument and the main clause demonstrative representing the object of *say*.

The Ugric facts surveyed above suggest a more plausible developmental path for the Germanic complementizer. The canonical correlative construction has been claimed to be the first type of finite subordination in the early Indo-European languages (Lehmann 1980), among them Germanic, as well (Kiparsky 1995; Harbert 2007: 422). In these languages, too, the relative pronoun was the first element with a subordinator role, the subordinating function of which was later extended to complement clauses. The most common, hence most likely context where clausal complementation could grammaticalize is the context of *say*-verbs crosslinguistically. *Say*-verbs were presumably transitive in Proto-Germanic, co-occurring with an accusative cataphoric pronoun, therefore when they integrated their propositional complements, their accusative pronoun evoked a nominal (rather than adverbial) relativizer as a subordinator – as attested in (46). In (46), the complement clause is still a DP-adjunct, but with the disappearance of the main clause pronoun it becomes a direct complement of V.

6. Summary

This paper has argued that the Ugric languages provide factual evidence for the discredited assumption that there is an evolutionary path from parataxis to finite subordination, as shown in (49):

(49) parataxis → correlative clause → finite relative clause
→ finite complement clause

The Ugric evidence also puts certain early Indo-European data in a new light, suggesting that they represent stages of a grammaticalization path similar to that documented in Uralic.

The main reason why the parataxis-to-hypotaxis hypothesis has lost credibility is the association of parataxis with a less advanced stage of human language. This paper has shown that parataxis fulfils a particular role in fully developed SOV languages with non-finite subordination. Non-finite subordinate clauses cannot have independent tense and do not have a force projection, hence predicates that select propositional complements with independent tense and force, primarily verbs of communication and propositional attitude, take paratactic complement clauses.

In the Ugric languages, the first finite subordinate clause type to emerge was the correlative construction. It has been argued that the correlative interpretation of this pattern is a recent development; originally, the construction was understood to involve two paratactically juxtaposed coordinate clauses, with an indefinite / indeterminate pronoun in the first clause. This finding converges with the theory of Belyaev and Haug (2020), who claim based on semantic evidence that the correlative clause derives from a paratactic sentence with a wh-pronoun used as an indefinite. A structural ambiguity similar to that attested in the Ugric languages has been pointed out in the Early Hittite (Hahn 1946; Motter 2023) and Early Latin correlative patterns (Kroll 1910), which suggests that a similar evolution from parataxis to the correlative construction happened in the Indo-European languages, as well.

In Hungarian, where a drift to SVO syntax started at least a thousand years earlier than in its Ob-Ugric sister languages, we could also follow the further evolution of the correlative construction. It has been shown to be the source of both finite relative clauses, and finite complement clauses. The path to finite relativization involved the reversal of the relative clause – main clause and order, and the reanalysis of the relative clause as the modifier of its main clause correlate. A similar evolution from the correlative clause to the adnominal relative clause has been assumed to have taken place in the Indo-European languages (see Ram Prasad 2022 and the literature cited by him).

Finite complementation evolved in the context of *say*-verbs. At the SOV stage of Hungarian, *say*-verbs took a paratactic propositional complement heralded by the proadverb *úgy* ‘so’ in the reporting sentence. When the loosening of SOV order allowed the incorporation of paratactic propositional complements into the reporting sentence, *úgy* ‘so’ evoked the use of *hogy* ‘how’, its *wh* correlate, as a subordinator. This move involved the reanalysis of *hogy* from a phrasal relative preposed into Spec,CP from a VP-adjoined position to a complementizer base-generated in C. The declarative complementizers derive from relative pronouns in several Indo-European languages, as well. The evolution of the German *dass*, English *that* has been traced back by Axel-Tober (2017) as far as the explicative construction, which represents an intermediate stage in the relative pronoun-to-complementizer development of Hungarian *hogy*. The Hungarian parallel helps to reconstruct the undocumented early phase of the grammaticalization of relative-based complementizers crosslinguistically.

References

Aikio, Ante. 2019. Proto-Uralic. *Oxford Guide to the Uralic Languages*, ed. by Marianne Bakró-Nagy, Johanna Laakso, and Elena Skribnik. Oxford: Oxford University Press.

Axel-Tober, Katrin. 2017. The development of the declarative complementizer in German. *Language* 93(2).e29–e65.
Bacskai-Atkari, Julia and Éva Dékány. 2014. From non-finite to finite subordination. The history of embedded clauses. *The evolution of functional left peripheries in Hungarian syntax*, ed. by Katalin É. Kiss, 148–223. Oxford: Oxford University Press.
Bacskai-Atkari, Julia and Éva Dékány. 2021. Cyclic changes in Hungarian relative clauses. *Syntactic Features and the Limits of Syntactic Change*, ed. by Jóhannes Gísli Jónsson, and Thórhallur Eythórsson, 40–63. Oxford: Oxford University Press.
Belyaev, Oleg, and Dag Haug. 2020. The genesis and typology of correlatives. *Language* 96. 874–907.
Bende-Farkas, Ágnes. 2014. From A-quantification to D-quantification: Universal quantifiers in the sentence and in the noun phrase. *The evolution of functional left peripheries in Hungarian syntax*, ed. by Katalin É. Kiss, 83–121. Oxford: Oxford University Press.
Bhatt, Rajesh 2003. Locality in correlatives. *Natural Language and Linguistic Theory* 21(3).485–541. https://www.jstor.org/stable/4048041
Csepregi, Márta. 2011. *Szurguti osztják chrestomathia*. Szeged: JATE.
Csepregi, Márta. 2012. Participiális jelzős szerkezetek két hanti nyelvjárásban [Participial attributive constructions in two Khanty dialects]. *Nyelvtudományi Közlemények* 108.61–94.
Csepregi, Márta. 2023. Khanty. In *The Uralic Languages*, edited by Daniel Abondolo and Riitta-Liisa Valijärvi, Ch. 16. Milton Park: Routledge.
Davidson, Donald. 1968. On saying 'that'. *Synthese* 130–46.
Dayal, Veneeta. 1996. *Locality in WH quantification: Questions and relative clauses in Hindi* (Studies in Linguistics and Philosophy 62). Boston: Kluwer.
Dékány, Éva, Katalin Gugán & Orsolya Tánczos. 2020. Contact-induced change in Surgut Khanty relative clauses. *Folia Linguistica* 54(1).1–43. http://doi.org/10.1515/flin-2020-2026
Delbrück, Berthold. 1900. *Vergleichende Syntax der indogermanischen Sprachen*. Vol. 3. Strasbourg: Karl J. Trübner.
Dömötör Adrienne. 2008. Az *úgy mond*-tól az *úgymond*-ig. Egy diskurzusjelölő elem története az ómagyar kortól napjainkig [From *úgy mond* to *úgymond*. The history of a discourse particle from Old Hungarian to the present]. *Magyar Nyelvőr* 132(1).37–52.
DEWOS = Steinitz, Wolfgang 1967–1991
É. Kiss, Katalin. 2023a. The (non-)finiteness of subordination correlates with basic word order: Evidence from Uralic. *Acta Linguistica Academica* 70(2).171–194. DOI: https://doi.org/10.1556/2062.2023.00647
É. Kiss, Katalin. 2023b. From relative proadverb to declarative complementizer: The evolution of the Hungarian *hogy* 'that'. *The Linguistic Review* 40(1).107–30. https://doi.org/10.1515/tlr-2022-2107
É. Kiss, Katalin, and Katalin Gugán. 2021. A *hogy* kötőszó kialakulása [The evolution of the complementizer *hogy*]. *Nyelvtudományi Közlemények* 117.199–226. DOI: 10.15776/NyK.2021.117.7
Filchenko, Andrey. 2010. *Aspects of the grammar of Eastern Khanty*. Tomsk: Tomsk State Pedagogical University.
Giannakidou, Anastasia. 2001. The meaning of free choice. *Linguistics and Philosophy* 24.659–735.
Grosu, Alexander, and Fred Landman. 1998. Strange relatives of the third kind. *Natural Language Semantics* 6(2).125–70. DOI: https://doi.org/10.1023/A:1008268401837
Gulya, János. 1966. *Eastern Ostyak chrestomathy* (Indiana University Publications Uralic and Altaic Series 51) . The Hague: Mouton.
Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and the composition of the left periphery*. Oxford: Oxford University Press.

Hahn, Adelaide. 1946. The origin of the relative *kwi-*, *kwo-*. *Language* 22(1).68–85.
Halm, Tamás. 2016. The syntactic position and quantificational force of FCIs in Hungarian. *Acta Linguistica Hungarica* 63(2).241–76.
Harbert, Wayne. 2007. *The Germanic Languages*. Cambridge: Cambridge University Press.
Harris, Alice C., and Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
Heine, Bernd, and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.
Hettrich, Heinrich. 1988. *Untersuchungen zur Hypotaxe im Vedischen* [Studies in Indo-European Language and Culture; New Series, 4]. Berlin and New York: Walter de Gruyter.
Hoffner, Harry A. and H. Craig Melchert 2008. *A grammar of the Hittite language*. Winona Lake, Indiana: Eisenbrauns.
Hopper, Paul J. and Elisasbeth Closs Traugott 2003. *Grammaticalization*. 2nd ed. Cambridge: Cambridge University Press.
Jahnhunen Juha. 1981. On the structure of Proto-Uralic. *Finnisch-ugrische Forschungen* 44.23–42. Helsinki: Société finno-ougrienne.
Justus, Carol. 1976 . Relativization and topicalization in Hittite. *Subject and topic*, ed. by Charles N. Li, 213–46. New York: Academic Press.
Keshet, Ezra. 2013. Focus on conditional conjunction. *Journal of Semantics* 30(2).211–256. DOI: https://doi.org/10.1093/jos/ffs011
Kiparsky, Paul. 1996. The shift to head-initial VP in Germanic. *Studies in comparative Germanic syntax*, ed. by Höskuldur Thräinsson, Samuel David Epstein, and Steve Peter, 140–79. Dordrecht: Kluwer.
Kratzer, Angelika, and Junko Shimoyama. 2002. Indeterminate pronouns: The view from Japanese. *Proceedings of Third Tokyo Psycholinguistics Conference*, ed. by Yukio Otsu, 1–25. Tokyo: Hituzi Syobo.
Kroll, Wilhelm. 1910. Der lateinische Relativsatz. *Glotta* 3(1).1-18.
Langacker, Ronald. 1969. On pronominalization and the chain of command. *Modern studies in English*, ed. by David A. Reibel and Sanford A. Schane, 160–86. Englewood Cliffs, NJ: Prentice-Hall.
Lehmann, Winfred P. 1980. The Reconstruction of Non-Simple Sentences in Proto-Indo-European. *Linguistic Reconstruction and Indo-European Syntax*, ed. by Paolo Ramat, 113–44. Amsterdam: John Benjamins.
Lipták, Anikó ed. 2009. *Correlatives cross-linguistically*. Amsterdam and Philadelphia: John Benjamins.
Lipták, Anikó. 2012. Correlative topicalization. *Acta Linguistica Hungarica* 59(3).245–302. DOI: 10.1556/ALing.59.2012.3.1
Motter, Thomas. 2023. Hittite correlatives are paratactic. *Glossa: a Journal of General Linguistics* 8(1). doi: https://doi.org/10.16995/glossa.8900
Munro, Pamela. 1982. On the transitivity of 'say' verbs. *Syntax and Semantics* 15.308–18.
Narrog, Heiko, and Bernd Heine. 2017. *Grammaticalization*. Oxford: Oxford University Press.
Németh, Endre and József Vígh. 2021. Párhuzamok és paradoxonok az uráli nyelvű népek populációgenetikai és nyelvi kapcsolatrendszereinek összehasonlítása során [Parallels and paradoxes in the comparison of the population genetic and linguistics links of Uralic speaking peoples]. *Anthropologiai Közlemények* 62.21-55.
Nikolaeva, Irina. 1999. *Ostyak* (Languages of the World/Materials 305). München: Lincom Europa.

Partee, Barbara. 1991. Topic, focus, and quantification. *Proceedings of SALT 1*, ed. by Steven K. Moore and Adam Zachary Wyner, 159–88. https://journals.linguisticsociety.org/proceedings/index.php/SALT/issue/view/136
Ram-Prasad, Krishnan J. 2022. The syntax of relative clauses and related phenomena in Proto-Indo-European. PhD dissertation, University of Cambridge.
Ravila, Paavo. 1960. Proto-Uralic. In Collinder, Björn. 1960. *Comparative Grammar of the Uralic Languages*, 250-51. Stockholm: Almqvist & Viksell.
Riese, Timothy. 2001. *Vogul* (Languages of the World/Materials 158). München: Lincom Europa.
Roberts, Ian. 2007. *Diachronic Syntax*. Oxford: Oxford University Press.
Srivastav, Veneeta 1991. The syntax and semantics of correlatives. *Natural Language and Linguistic Theory* 9.637–86.
Steinitz, Wolfgang. 1967–1991. *Dialektologisches und etymologisches Wörterbuch der ostjakischen Sprache*. 1–14. Lieferungen. Berlin: Akademie Verlag.
Szeverényi, Sándor and Katalin Sipőcz. 2019. Szibériai uráli nyelvek idéző szerkezetei [Reportative constructions of Uralic languages in Siberia]. In *Nyelvelmélet és kontaktológia 4*, ed. by Katalin É. Kiss, Attila Hegedűs, and Lilla Pinter, 127–46. Budapest – Piliscsaba: PPKE BTK. https://btk.ppke.hu/karunkrol/intezetek-tanszekek/angol-amerikai-intezet/elmeleti-nyelveszet-tanszek/tanszeki-kiadvanyok/nyelvelmelet-es-kontaktologia-4-2019
Türk, Attila. 2023. History. The archaeological background of Hungarian prehistory. In *Psalmus Hungaricus. A Hungarian cultural history I. From the beginnings to 1526*, edited by József Szentpétery, 16–20. Budapest: Bölcsészettudományi Kutatóközpont.
Walkden, George. 2018. Proto-Indo-European: A language without merge? Talk presented at the Workshop on Syntax and Reconstruction, University of York. http://www.sfs.uni-tuebingen.de/fileadmin/user_upload/AS/WalkdenOSWS1819.pdf
Weiss, Helmut. 2020. Where do complementizers come from and how did they come about? *Evolutionary Linguistic Theory* 2(1).30–55. https://doi.org/10.1075/elt.00014.wei

Primary sources

Apor Codex. Original: after 1416; Copy: end of the 15th century. György Volf (ed.) *Apor-kódex*. Budapest, 1879.
Bécsi Codex. Original: after 1416; Copy: around 1450. Gedeon Mészöly (ed.) *Bécsi kódex*. Budapest: MTA, 1916.
Birk Codex. 1474. István Pusztai (ed.) *Birk-kódex*. Budapest: Akadémiai Kiadó, 1960.
Bod Codex. Beginning of the 16th century. István Pusztai (ed.) *Bod-kódex*. Budapest: Magyar Nyelvtudományi Társaság, 1987.
Černecov, Valerij Nikolaevič. 1935. OUDB Northern Mansi Corpus. Text ID 1234, 1235, ed. by Gwen Eva Janda. http://www.oudb.gwi.uni-muenchen.de/?cit=1234
Cornides Codex. 1514–1519. András Bognár and Ferenc Levárdy (eds.) Cornides-kódex. Budapest: Akadémiai Kiadó, 1967.
Érdy Codex. 1524–27. Transcribed by the Sermones project, Budapest: ELTE. http://omagyarkorpusz.nytud.hu/documents/kodexek/ErdyK/3a/ErdyK_3a.txt
Érsekújvári Codex. 1529–1531. Haader Lea (ed.) *Érsekújvári kódex*. Budapest: Tinta Könyvkiadó, 2013.
Festetics Codex. Before 1494. Csilla N. Abaffy (ed.) *Festetich-kódex*. Budapest: Argumentum Kiadó.
Halotti beszéd [Funeral Sermon] 1195. *Régi magyar nyelvemlékek* [Old Hungarian linguistic documents], ed. by Adrienne Dömötör, p. 27. Budapest: Akadémiai Kiadó.

Jogi szabály [Legal rule] 1476–1490. *Régi magyar nyelvemlékek* [Old Hungarian linguistic documents], ed. by Adrienne Dömötör, p. 46. Budapest: Akadémiai Kiadó.
Jókai Codex. Original: after 1372; Copy: around 1448. P. János Balázs (ed.) *Jókai-kódex.* Budapest: Akadémiai Kiadó, 1981.
Jordánszky Codex. 1516–1519. Ferenc Toldy and György Volf (eds.) *A Jordánszky kódex bibliafordítása*. Buda, 1888.
Káldi, György (transl.). 1626. *Szent Biblia*. Wien.
Kannisto, Juha Artturi. 1951–1955. Wogulische Volksdichtung I–II, ed. by Matti Liimola. Helsinki: Suomalais-Ugrilainen Seura.
Károli, Gáspár (transl.). 1590. *Szent Biblia*. Vizsoly. Facsimile published by Helikon Kiadó, Budapest, 1990.
Königsbergi töredék [Königsberg Fragment]. Middle of 14th century. *Régi magyar nyelvemlékek* [Old Hungarian linguistic documents], ed. by Adrienne Dömötör, p. 32. Budapest: Akadémiai Kiadó.
Könyvecse. 1521. István Pusztai (ed.) *Könyvecse az szent apostoloknak méltóságokról* [Booklet on the dignities of the saint apostles]. Budapest: Magyar Nyelvtudományi Társaság, 1986.
Kulcsár Codex. 1539. Lea Haader and Zsuzsanna Papp (eds.) *Kulcsár-kódex.* Budapest: MTA Nyelvtudományi Intézet, 1999.
Lobkowicz Codex. 1514. Andrea Reményi (ed.) *Lobkowicz-kódex*. Budapest: Argumentum Kiadó.
Margit legenda [Margaret legend]. 1510. János P. Balázs, Adrienne Dömötör, and Katalin Gugán (eds.) *Margit-legenda*. Budapest: Magyar Nyelvtudományi Társaság, 1990.
Munkácsi, Bernát. 1896. *Vogul népköltési gyűjtemény* [Vogul folklore collection] IV. Budapest: MTA.
Müncheni Codex. Original: after 1416; Copy: 1466. Antal Nyíri (ed.) A Müncheni kódex 1466-ból. Budapest: Akadémiai Kiadó, 1971.
Old Hungarian Database http://oldhungariancorpus.nytud.hu/
OUDB Ob-Ugric Database http://www.babel.gwi.uni-muenchen.de/index.php?abfrage=welcome_eurobabelandnavi=aboutandsubnavi=
Paasonen, Heikki. 1901. Surgutostjakische Textsammlungen am Jugan. Published as Edit Vértes (ed.) H. *Paasonens surgutostjakische Textsammlungen am Jugan* (Mémoires de la Société Finno-ougrienne 240). Helsinki: Suomalais-Ugrilainen Seura, 2001.
Példák könyve [Book of parables]. 1510. András Bognár and Ferenc Levárdy (eds.) *Példák könyve*. Budapest: Akadémiai Kiadó, 1960.
Pesti Mizsér, Gábor (transl.). 1536. *Új Testamentum magyar nyelven*. Viennae Pannonie. Facsimile edition: Budapest: MTA, 2002.
Sylvester, János (transl.). 1541. *Új Testamentum*. Újsziget: Abádi Benedek. Facsimile ed. by Béla Varjas. Budapest: Akadémiai Kiadó, 1960.