The genesis and typology of correlatives: Supplementary materials

Oleg Belyaev, Dag Haug

Language, Volume 96, Number 4, December 2020, (Article)

Published by Linguistic Society of America

DOI: https://doi.org/10.1353/lan.2020.0083

For additional information about this article
https://muse.jhu.edu/article/776463/summary

For content related to this article
https://muse.jhu.edu/related_content?type=article&id=776463
The genesis and typology of correlatives

Typological dataset

Oleg Belyaev and Dag Haug
Lomonosov Moscow State University and University of Oslo

Definition and sample

This document contains the relevant information about correlative constructions in our sample. The sample includes virtually all constructions for which sufficient information is available and which conform to our definition of correlatives as given in section 2.1 of the main paper, repeated here for convenience:

(1) A correlative structure consists of two clauses, the correlative clause and the matrix clause, where

1. a noun phrase containing a relativizer occurs in the correlative clause;
2. a noun phrase anaphorically related to the first one occurs in the matrix clause;
3. the correlative clause is wholly to the left of the matrix clause.

Note that, due to requirement of an NP-internal relativizer and the preposition of the relative clause, this definition excludes certain constructions that have sometimes been labeled as ‘correlatives’. Two examples are particularly noteworthy. Australian ‘adjoined clauses’ as in (2) can mark any kind of subordinate clause and have no marking apart from a general subordinating morpheme on the verb. The tradition since K. L. Hale (1976) is to treat such constructions as a special kind of subordinate clause, but others, such as Bittner (2001), have argued that they are, in fact, structurally and semantically analogous to correlatives. Since no special marking is found on the relative NP (which may even be absent altogether, as in ex. 2), and the subordinate clause is postposed, these do not fit our definition; in fact, this result is justified as there is no consensus in the literature on the extent to which Australian ‘adjoined clauses’ are akin to correlatives (see Nordlinger 2006).

(2) Warlpiri (Pama-Nyungan)

Ngajulu-rlu rna yankirri pantu-rnu, [kuja-lpa ngapa nga-rnu ]

I-ERG AUX emu spear-PST COMP-AUX water drink-PST

‘I speared the emu which was / while it was drinking water.’ (K. L. Hale 1976, 78)

Another similar case is Modern Georgian, where a new construction (3) developed that some authors, e.g. Hendery (2012, 180ff.), treat as a correlative. In these construction, the subordinate clause precedes the main clause, which contains a noun phrase containing a demonstrative; the subordinate clause is introduced by the preverbal universal subordinator ro(m)² which is also used for many other types of subordinate clauses (see Hewitt 1987 for an overview of the Georgian system of subordination).

(3) Georgian (South Caucasian)

[a-m k’ac-ma rom naxa ], i-m kal-s vicnob

this-DET.OBL man-ERG COMP he.saw.her.AOR that-DET.OBL woman-DAT I.know.her

‘I know that woman which this man saw.’ (Hewitt 1987, 188)

The relative NP, when present, is often accompanied by a demonstrative, as in (3), which may make it seem that this construction is a dem-correlative; however, this is not necessarily so: the clause-internal NP can be left unmarked (4) or be absent altogether (5).

(4) Georgian (South Caucasian)

[Durmišxan-s Alget-ze rom c’iskvili eč’ira ], i-s c’iskvili...

Durmišxan-DAT Alget-on COMP mill he.have.it that-DET.NOM mill

‘the mill which Durmishxan had on Alget …’

(Harris 1994, 134 apud Hendery 2012, 180)

(5) Georgian (South Caucasian)

[šen-gan ro miviɣeb ], i-m pul-it me gadavixdi val-s

you-from COMP I.receive.it that-DET.OBL money-INS I.NOM I.pay.it debt-DAT

‘I will pay off the debt with that money which I receive from you.’

(Harris 1994, 133 apud Hendery 2012, 180)

Thus, Georgian conforms to all the requirements in (1) but one: it does not use a designated relativizer. In fact, other linguists, such as Harris (2000), treat this construction as a ‘prenominal gapped relative’ rather than a true correlative; if this is the case, the examples with an overt ‘internal head’ are instances of resumption, not unheard of in prenominal gapped relative clauses in other languages; a well-studied example is Japanese, see e.g. McCawley (1976).

Thus, the status of Georgian prenominal relative clauses must be clarified before they can be included in any typology of correlatives.²

² The glossing of ro(m) has been changed from ‘that’ to ‘comp’ to avoid ambiguity: this marker is a complementizer, not a demonstrative.

4. It should be noted, however, that, if anything, Georgian prenominal relative clauses are dem-correlatives; if so, they seem to conform to our proposal, as all examples in the literature that we are aware of involve the definite reading.
More generally, it is not clear how ‘correlative-like’ constructions with no overt obligatory marking of the relative NP, such as in Australian languages or possibly in Georgian, should be treated: they seem intuitively similar to correlatives but significantly different at the same time. It may be that, while sharing certain structural and semantic properties, they ultimately involve a different scenario of grammaticalization. Note that both Warlbiri and Georgian use universal subordinators or complementizers (kuja and rom, respectively), which is not attested in any language in our sample, which either use no overt clause-level marker or a specialized marker, such as a conditional morpheme or an indirect question particle.

Features and their values

The values of the columns are as follows: type refers to the etymology of the relativizer: wh means that the relativizer is formally identical to an interrogative (which in most cases can be shown to also be used as an indefinite), wh-indef means that it is an indefinite pronoun which is morphologically derived from an interrogative one, and dem means that it is demonstrative. univ and def refer to the availability of the universal and definite reading (see section 2.2 of the main paper): + means it is available, +/- that it is only available with special marking, and – that it is unavable. +/- was only used in cases where the special marking is minimal and additive: specifically the conditional marker, the subjunctive mood or generic tense morphology is involved in the cases at hand, as discussed in the description of the relevant languages. Whenever there are more substantial differences, we postulated two different correlative constructions instead. (This happens in Tamil and Dargwa.) [i=i] is adopted from Gärtner (2009) and refers to the indefinite-interrogative ambiguity (see section 4.1 of the main paper): % means that interrogatives can be used with indefinite meaning in conditional structures (and possibly some other embedded structures) and – that interrogatives cannot be used as indefinites. cond. refers to the presence of (relics of) conditional morphosyntax in correlative constructions. The latter two columns are only relevant for wh-based correlatives, so dem-based correlatives have n/a in these columns.

<table>
<thead>
<tr>
<th>language</th>
<th>type</th>
<th>univ.</th>
<th>def.</th>
<th>[i=i]</th>
<th>cond.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indo-European</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hittite (§ 1.1)</td>
<td>wh</td>
<td>+</td>
<td>+</td>
<td>%</td>
<td>–</td>
</tr>
<tr>
<td>Latin (§ 1.2)</td>
<td>wh</td>
<td>+</td>
<td>+</td>
<td>%</td>
<td>–</td>
</tr>
<tr>
<td>Ancient Greek (§ 1.3)</td>
<td>dem</td>
<td>+/–</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Old Czech, dem (§ 1.4.1)</td>
<td>dem</td>
<td>–</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Old Czech, wh (§ 1.4.1)</td>
<td>wh</td>
<td>+</td>
<td>–</td>
<td>%?</td>
<td>–</td>
</tr>
<tr>
<td>Serbo-Croatian (§ 1.4.2)</td>
<td>wh</td>
<td>+</td>
<td>–</td>
<td>%</td>
<td>–</td>
</tr>
<tr>
<td>Russian (§ 1.4.3)</td>
<td>wh</td>
<td>+</td>
<td>–</td>
<td>%</td>
<td>–</td>
</tr>
<tr>
<td>Ossetic, simple (§ 1.5.1)</td>
<td>wh</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ossetic, multiple (§ 1.5.1)</td>
<td>wh</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Old Indic (§ 1.6.1)</td>
<td>dem</td>
<td>+/–</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Hindi (§ 1.6.2)</td>
<td>dem</td>
<td>+/–</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Kalasha (§ 1.6.3)</td>
<td>wh</td>
<td>+</td>
<td>+</td>
<td>%?</td>
<td>–</td>
</tr>
<tr>
<td>Shina (§ 1.6.4)</td>
<td>wh</td>
<td>+</td>
<td>+</td>
<td>%</td>
<td>–</td>
</tr>
</tbody>
</table>

5. Following Gärtner (2009) we would use + for languages where interrogatives can be used as indefinites in all contexts, but this does not occur in our sample.
### Table 1: (continued)

<table>
<thead>
<tr>
<th>language</th>
<th>type</th>
<th>univ.</th>
<th>def.</th>
<th>[i=i]</th>
<th>cond.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Uralic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungarian ($\S\ 2.1$)</td>
<td>wh</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Old Hungarian ($\S\ 2.1$)</td>
<td>wh-indef</td>
<td>+</td>
<td>−</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Udmurt ($\S\ 2.2$)</td>
<td>wh</td>
<td>+</td>
<td>+</td>
<td>%</td>
<td>−</td>
</tr>
<tr>
<td><strong>Turkic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Old Turkic ($\S\ 3.1$)</td>
<td>wh</td>
<td>+</td>
<td>−</td>
<td>%?</td>
<td>+</td>
</tr>
<tr>
<td><strong>Tungusic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Udihe$_{\text{wh}}$ ($\S\ 4.1$)</td>
<td>wh</td>
<td>+</td>
<td>−</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Udihe$_{\text{dem}}$ ($\S\ 4.1$)</td>
<td>dem</td>
<td>−?</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>Yenisean</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ket$_{\text{wh}}$ ($\S\ 5$)</td>
<td>wh</td>
<td>+</td>
<td>−?</td>
<td>?</td>
<td>−</td>
</tr>
<tr>
<td>Ket$_{\text{dem}}$ ($\S\ 5$)</td>
<td>dem</td>
<td>−?</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td><strong>Basque</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basque ($\S\ 6$)</td>
<td>wh-indef</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td><strong>Dravidian</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tamil$_{\text{COND}}$ ($\S\ 7.1$)</td>
<td>wh</td>
<td>+</td>
<td>−</td>
<td>%</td>
<td>+</td>
</tr>
<tr>
<td>Tamil$_{\text{FIN}}$ ($\S\ 7.1$)</td>
<td>wh</td>
<td>−</td>
<td>+</td>
<td>%</td>
<td>−</td>
</tr>
<tr>
<td>Old Tamil ($\S\ 7.2$)</td>
<td>wh</td>
<td>+</td>
<td>−</td>
<td>%</td>
<td>+</td>
</tr>
<tr>
<td><strong>East Caucasian</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kryz ($\S\ 8.1$)</td>
<td>wh</td>
<td>+</td>
<td>−?</td>
<td>%</td>
<td>+</td>
</tr>
<tr>
<td>Khwarshi ($\S\ 8.2$)</td>
<td>wh</td>
<td>+</td>
<td>−?</td>
<td>%?</td>
<td>−</td>
</tr>
<tr>
<td>Tsez ($\S\ 8.3$)</td>
<td>wh</td>
<td>+</td>
<td>−?</td>
<td>%?</td>
<td>−</td>
</tr>
<tr>
<td>Dargwa$_{\text{GEN}}$ ($\S\ 8.4$)</td>
<td>wh</td>
<td>+</td>
<td>−</td>
<td>%</td>
<td>+</td>
</tr>
<tr>
<td>Dargwa$_{\text{SPEC}}$ ($\S\ 8.4$)</td>
<td>wh</td>
<td>−</td>
<td>+</td>
<td>%</td>
<td>−</td>
</tr>
<tr>
<td><strong>Sino-Tibetan</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese ($\S\ 9.1$)</td>
<td>wh</td>
<td>+</td>
<td>−</td>
<td>%</td>
<td>+</td>
</tr>
<tr>
<td>Tibetan ($\S\ 9.2$)</td>
<td>wh</td>
<td>+</td>
<td>+</td>
<td>%?</td>
<td>+</td>
</tr>
<tr>
<td><strong>Niger-Congo</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kita ($\S\ 10.1$)</td>
<td>dem</td>
<td>+/−</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Mandinka ($\S\ 10.1$)</td>
<td>dem</td>
<td>+/−</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Bambara ($\S\ 10.1$)</td>
<td>dem</td>
<td>−</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Soninke ($\S\ 10.1$)</td>
<td>dem</td>
<td>−</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Vai ($\S\ 10.1$)</td>
<td>dem</td>
<td>−?</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>Supyire ($\S\ 10.2$)</td>
<td>dem</td>
<td>+/−</td>
<td>+</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>
1 Indo-European

1.1 Anatolian: Hittite [WH]

Readings Correlatives were the main strategy of relativization in Hittite. They had both the definite and universal readings, which are referred to in the literature as *determinate* and *indeterminate*, respectively (Held 1957; M. Hale 1987, 46–49; Garrett 1994, 43–49; Hoffner and Melchert 2008, 424–425). The following examples illustrate both types:

(6) KUR.KUR.ḪI.A kue dannatta ammuk EGIR-pa ašešanun #
    land.ACC.PL.N REL:ACC.PL.N empty:ADJ:ACC.PL NOM.SG.C RE.PV settle:1SG.PRET.ACT
    nu=mu=kan apē=ya ḫūmanda arḥa dāš
    CONN=CL.1SG.DAT=PTCL.LOC this:ACC.PL.N all:ACC.PL.N away.PV take-3SG.PRET.ACT

‘The empty lands which I resettled, all those he took away from me.’ (Yates 2014, 2)

(7) nu=šši=ššan kuit šaḫḫan LUGAL-uš
    CONJ=CL.3SG.DAT=PTCL.LOC REL:INANIM:ACC.SG.N SERVICE:ACC.SG.N KING-NOM.SG.C
    dāi # nu apāt ėššai
    place:3SG.PRS.ACT CONN that:ACC.SG.N do:3SG.PRS.ACT

‘Whatever service the king imposes on him, (he) will do that.’ (Yates 2014, 2)

Unfortunately, there are no good examples of quantification in episodic contexts in the literature; most of the so-called indeterminate contexts involve genericity or indefinite future-tense reference. However, it is well-established that the two types of correlative readings have different syntactic properties in Hittite. Thus it can be safely assumed that these are two genuine readings, and the generic reading is not derived from the definite one.

Interrogative-indefinite affinity The indefinite pronoun in Hittite is formed from the interrogative via suffixation: kuiš ‘who’ → kuiški ‘someone’ (Hoffner and Melchert 2008, 149). ‘Only in combination with takku/mān ‘if’ can kui- have an indefinite meaning’ (Hoffner and Melchert 2008, 149). This usage is seen in the following example:

(8) nu=kan mān AWAT NARARI kuwapi šarā išparzazi #
    CONN=LOC:PTCL if summons assistance when up come:PRS.3SG

‘If at some point (lit. If when) a summons for assistance comes up…’

(Hittite, NH/INS (CTH 106.A.1) Bo 86/299 rev. iii 44–46; Andrej Sideltsev, p.c.)

1.2 Italic: Latin [WH]

Readings In Latin, *wh*-based correlative structures are quite common in the early texts, and we find both universal and definite readings (9)–(10).

---

6. We are grateful to Andrej Sideltsev for drawing our attention to these examples. The glossing in Hittite examples deviates somewhat from the Leipzig standard: the hyphen (-) separates Akkadograms and Sumerograms from transcribed Hittite forms, while the period sign (.) separates several Sumerograms and Akkadograms from each other.
9. [cui testimonium defuerit], is tertius diebus
   REL.DAT testimony:NOM.SG is.absent:FUT.PRF.3SG, he:NOM third:ABL.PL day:ABL.PL
   ob portum obuagulatum ito.
   to door:ACC.SG to.summon go:IMP.3SG
   'he whose witness is absent, he shall go to summon him every third day'
   (The twelve tables II.3, Warmington)

10. [quam earum in iis locis optimam dicent
    REL.ACC.SG them:GEN.PL in these:ABL.PL places:ABL.PL best:ACC say:FUT.3PL
    esse ], eam maxime serito
    be:INF, that:ACC in.particular sow:IMP.2SG.FUT
    '(the one) which they say is best in these places, sow that one in particular'
    (Cato, de Agricultura 6.1)

As pointed out by Sturtevant (1930, 148), the universal reading is much more frequent than
the definite one in early Latin. It is, however, not clear whether this reflects a primacy for the
generalizing meaning, or whether it is a a genre effect, see e.g. the discussion in Probert and

Interrogative-indefinite affinity  Latin is included in the list of [%i=i] languages in Gärtner
(2009, 13).

1.3 Greek [DEM]

Ancient Greek displays a wide array of strategies for relativization from the earliest stages.⁷
Alongside postnominal relative clauses and free relative clauses, we also find correlatives. The
pronoun hos used in relative clauses of all types is traditionally taken to be a demonstrative in
origin.

Both correlatives and free relatives display the expected maximalizing semantics yielding a
definite reading (11). However, correlatives and free relatives in the indicative are not normally
used for universal readings — this meaning requires the use of the subjunctive mood and (in
most dialects) a modal particle ke or an (12).

11. hos de moi oios eːn (...) ton su proːen
    REL.NOM but me:DAT alone:NOM be:PST.3SG him:ACC you:NOM just.now
    kteinas
    killed:PFV.PST.2SG
    'He that alone was left me, him you just now killed.' (Homer, Iliad 24.499–500)

12. hos ke τʰeois epipeitʰ:eːtai, mala t’ ekluon
    REL.NOM modal god:DAT.PL obey:SBJV.PRS.3SG, greatly and listen:PFV.PST.3PL
    autou
    him:GEN
    'Whoever obeys the gods, they listen greatly to him.' (Homer, Iliad 1.218)

⁷ We ignore Mycenaean Greek although there may be relative clauses in the material, see Probert (2015, 199-202).
Occasionally, maximalizing relatives can get a universal reading even when the verb is in the indicative (13)

(13) apʰretor eʰemistos anestios estin ekeinos, hos polemou
clanless:NOM lawless:NOM hearthless:NOM be:PRS.3SG he:NOM, REL:NOM war:GEN
eratai epideːmiou kruoentos
loves:PRS.3SG civil:GEN horrible:GEN
‘Clanless, lawless and hearthless is he who loves horrible civil war.’
(Homer, Iliad 9.63–64)

This construction is much rarer than the subjunctive construction; moreover, it always involves individual-level predicates in the relative clause (statives, or derived habituals, see Probert [2015], 92–93). As such, the construction generalizes over individuals rather than situations and may be akin to the generic use of definites in sentences like The lion has a mane.

1.4 Slavic

1.4.1 Old Czech [WH/DEM]

Readings Czech represents a unique case, as not only did it extensively employ correlatives, but correlatives with both interrogative and demonstrative-based pronouns have coexisted for some time in the history of the language. Their semantic distribution is quite remarkable. As Bauer ([1972], 40-41) demonstrates, in Old Czech ‘les [propositions relatives avec jenž, ješto]’ (i.e. yo-based relatives — O.B. & D.H.) indiquaient d’habitude le caractère d’une personne ou d’une chose connue, les [propositions relatives avec kto, čso, ktery]’ (i.e. interrogative-based relatives — O.B. & D.H.) se rattachaient à une personne ou à une chose en général’. He provides the following examples to illustrate this point:

(14) Old Czech (Slavic > Indo-European), dem-based [DEM/WH]
    jehož někdy najvěrnějšieho mniše, tomu
    REL:MG.SG.ACC sometimes most.loyal:MG.SG.ACC consider:AOR.3PL that:MG.SG.DAT
    káza na nátonu bradu obrubati
    order:AOR.3SG on log:SG.LOC beard:SG.ACC cut:OFF:INF
    ‘He ordered to cut off the beard on the (wood) log to the one who was considered the most loyal.’ (Bauer [1972], adapted from The Chronicle of Dalimil 59,50)

(15) Old Czech (Slavic > Indo-European), wh-based [DEM/WH]
    kdo chce zbytí strasti, nerod’ prsta
    mezi dřvi a podvoj klásti
    between door:SG.INS and frame:SG.INS put:INF
    ‘Whoever wants to stop sorrows should not put two fingers between a door and its frame.’ (Bauer [1972], adapted from the Chronicle of Dalimil 85,58)

According to Bauer, later in the history of Czech, interrogative-based relative clauses gradually extended their function, such that they started to be used in both contexts, and jenž ‘which’

8. I am grateful to Maria Kholodilova for drawing my attention to these data.
and other dem-based correlatives were displaced to their marked stylistic (archaic) variants. This perfectly corresponds to our theory.

**Interrogative-indefinite affinity** We are not aware of any information on the licensing of indefinite readings of interrogatives in Old Czech. However, all other Slavic languages in our sample and in Gärtner (2009) are [%i=i], which makes this status a plausible assumption for Old Czech as well. In addition, modern Czech licenses indefinite readings of wh-pronouns at least in some modal constructions (Šimík 2009).

### 1.4.2 Serbo-Croatian [WH]

**Readings** Only the universal interpretation is available. For example, Arsenijević (2009) has proposed an analysis equating correlatives with conditional sentences based on Serbo-Croatian data.

(16) \[ Ko ima magarca , taj ga i bije. \]

- *REL.ANIM* has donkey that *CL.ACC* and beats
- 'Anyone who has a donkey beats it.’ (Arsenijević 2009, 142)

One of the arguments that Arsenijević uses to support his analysis, and one which is quite important alongside other languages in our sample, is that the conditional subordinator can optionally be inserted into the relative clause:

(17) \[ Ko tebe ako nečim udari, ti tome vrati istom \]

- *who you.ACC* if *something.INST* hits *you* that*one.DAT* return*IMP* same*INS* measure*INS* hit*IMP* in*DAT* same*INS* go*IMP* home*IMP*
- 'If anyone hits you with something, hit him in the same way as he hit you.'/ ‘Whoever hits you with something, hit him in the same way as he hit you.’

(Arsenijević 2009, 141)

**Interrogative-indefinite affinity** Serbo-Croatian is explicitly described as [%i=i] by Arsenijević (2009, 144):

(18) \[ Ako tebe ko nečim udari, ti idi kući. \]

- if *you.ACC* who *something.INS* hits *you* go*IMP* home*IMP*
- 'If anyone hits you with something, go home.'

(Arsenijević 2009, 144)

### 1.4.3 Modern Russian and other Slavic [WH]

**Readings** With varying degrees of explicitness, the same situation as in Serbo-Croatian is described for other modern Slavic languages. For example, as Nikunlasi (2008, 108) observes, correlatives in modern Standard Russian always involve universal quantification and can be paraphrased through conditional constructions. Similar parallels between conditionals and correlatives in Slavic are observed in Izvorski (1996) for Bulgarian.

Middle Russian used to also have a different construction, which used the pronoun *kotoryj* ‘which’ with an optional internal head; this construction is still marginally possible in the
modern colloquial language (Kholodilova 2014, 5.1.3). These could have definite interpretations:

(19) A [kotoraja gsdr’ lošed poslanaja s nim ]… i and REL.NOM.SG.F master horse.NOM.SG send.PTCP.NOM.SG.F with he.INS and ta lošed stala v Volodimire that.NOM.SG.F horse.NOM.SG stay.PST.SG.F in Vladimir.LOC.SG

‘As for the horse that was sent with Stephan, that horse stayed in the city of Vladimir.’ (Mitrenina 2010, 62)

However, Mitrenina (2010) argues that these constructions should more properly be called ‘pseudo-correlatives’, and the kotoryj is actually an indefinite pronoun or determiner rather than a relativization marker; the ‘correlate’ in the main clause is bound according to general rules of anaphora. According to this analysis, then, this Middle Russian construction is excluded from our sample.

Interrogative-indefinite affinity Russian is included in the list of [%i=i] languages in Gärtner (2009, 13).

1.5 Iranian

1.5.1 Ossetic [WH]

Readings In Ossetic, simple correlatives can have the universal and definite readings (Belyaev 2014a):

(20) [didinǯ-ətɜ čəžg-ən ba-lɜvar kotː-aj ], fetː-on wəj fæd-ə
flower-PL REL girl-DAT PV-present do-PST.2SG see.PFV-1SG that.GEN father-GEN
‘I saw the father of the girl who you gave flowers to.’ (Belyaev 2014a, 276–277)

(21) [uroč-ə či n3 wəd-i ], wəj direktor-mɜ a-s3w-3d
lesson-IN REL.ANIM.NOM NEG be-PST.3SG that.DEM headmaster-ALL PV-GO-IMP.3SG
‘Let whoever didn’t come to the lesson go to the headmaster.’ (Belyaev 2014b, 199)

However, Ossetic also has multiple correlatives, which are a separate construction with a different set of properties. The main semantic difference between simple and multiple correlatives is that the latter may only have a quantificational reading:

(22) [k3m-ən sə čəžg j3= žərd3-m3 fə-səd-iš ], wəj-ə ra-kʷərtː-a
who-DAT REL girl his heart-ALL PV-GO-PST.3SG that.DEM GEN PV-ASK-PST.3SG
‘Everyone married whichever girl he liked’, lit. ‘Who liked which girl, (he) married her.’ (Belyaev 2014b, 201)

This sentence cannot mean that a specific young man like a specific girl and married her; it can only have a quantificational interpretation. Other languages that have multiple correlatives, such as Hindi (Srivastav 1991, 670), do allow specific interpretations.

Unlike simple correlatives, where the relative markers/subordinators have already diverged from interrogatives proper and form a separate lexical class, pronouns used in multiple
correlatives are *bona fide* interrogatives; in particular, \(k^\text{w}o\) ‘when’ can be used in simple correlatives, but not in questions or multiple correlatives, while \(k\text{sd} \) ‘when’, conversely, can be used in questions and multiple correlatives, but not in simple correlatives. Multiple correlatives also do not have such a strict demonstrative requirement as simple correlatives; for example, in (22), only the direct object is overtly expressed by a demonstrative, the subject, coreferent with \(\text{çî} \) ‘who’, is pro-dropped. Thus the construction where only a generalizing reading is available is both more formally connected to interrogatives and closer in its syntax to standard discourse anaphora.

**Interrogative-indefinite affinity** Ossetic does not allow using interrogative as indefinites in embedded contexts. However, this can in part be explained by syntactic reasons: interrogatives and the majority of subordinators (including conditional \(k^\text{w}o\)) compete for the preverbal position (Erschler [2012], Belyaev [2014a]). The fact that Ossetic has a series of indefinite pronouns based on the combination of negation \(ni\)- and interrogatives suggests that it used to license indefinite readings of interrogatives in negative contexts at some point in the past.

### 1.6 Indo-Aryan

#### 1.6.1 Old Indic [DEM]

The situation in the oldest stage of Indic – Vedic Sanskrit – is similar to that of Greek, which is not unexpected since the relativizers of both languages have the same etymon, \(^*\text{yo-}\). Ordinary indicative correlatives mostly have a definite reading.

(23) sóma \(\text{[yās te mayobhūva ūtāyah sánti dāśúse \(\text{tabhīr no avitā bhava}\)}\)

\(\text{soma:voc REL.nom.pl.f yours delightful:nom.pl.f means.of.help:nom.pl.f be:prs.3pl}
\(\text{dāśúse \(\text{tabhīr no avitā bhava}\)}\)

\(\text{worshipper:dat those:ins.pl us helper:nom be:imp}\)

‘Soma, which delightful means of help are yours for the worshipper, with those be our helper.’ (RV 1.91.9)

The universal reading often has special marking in the form of a subjunctive.

(24) \(\text{[yās \(\text{tūbhīyam dāśan nā tām āmho aśnavat}\)}\)

\(\text{rel.nom.sg you.dat honour:prs.sbjv.3sg not that:acc fear reach:prs.sbjv.3sg}\)

‘Whoever honours you, fear shall not reach him.’ (RV 2.23.4b)

Like in Ancient Greek, indicative correlatives also sometimes have a universal reading. In the absence of a detailed semantic study like that of Probert (2015) on Greek, it is unclear whether these have a special meaning. Taken together, the state of affairs in the two oldest IE language to preserve the \(^*\text{yo-}\) relativizer suggests that the core meaning of such constructions is that the definite reading is the original state of affair. This was in fact the view of Sturtevant (1930), who argued that \(^*\text{k\text{"i} / *k\text{`o}}\) was used in generalizing relative clauses and \text{yo-} in definite ones (Sturtevant 1930, 148). We should note that the originally demonstrative function of \(^*\text{yo-}\) is not beyond doubt, although it is the common view.
1.6.2 Hindi [DEM]

Hindi, as most other modern Indo-Aryan languages, employs correlatives as the main, if not the only, strategy of relativization. Correlative structures permit both definite and universal readings. Only the definite reading is available in (25a), while (25b) has the universal reading due to the presence of an overt habitual operator. In Hindi, this is a strict requirement (Dayal 1995, 189).

(25)

a. \[ \text{jo laRki tez hai} \text{ vo aksar safal hotii hai} \]
\[ \text{REL girl smart is she often successful HAB is} \]
\[ \text{‘The girl who is intelligent is often successful.’ (unique individual)} \]

b. \[ \text{jo laRki tez hotii hai, vo aksar safal hotii hai} \]
\[ \text{REL girl smart HAB is she often intelligent HAB is} \]
\[ \text{‘A girl who is smart is often intelligent.’ (generic)} \] (Dayal 1995, 189)

The relative pronoun, of course, has the same probably demonstrative origin as the Old Indic yo-series.

1.6.3 Kalasha [WH]

Readings Kalasha (Dardic) has a correlative construction which Bashir (1988, 358–359) calls the ‘incipient relative-correlative’ because it uses interrogatives instead of a dedicated series of relative pronouns. This makes Kalasha almost unique among Indo-Aryan languages in employing a wh-based relativization strategy. While Bashir does not discuss the semantics of correlatives, her examples shows that both readings are available:

(26) \[ \text{ku’ra moc ki sarak’-una kas’-iu-dai} \]
\[ \text{REL ANIM man COMP road-LOC walk(PRS/FUT-ST)-3SG he my maternal uncle} \]
\[ \text{‘The man who is walking in the road is my maternal uncle.’} \] (Bashir 1988, 360)

(27) \[ \text{kiy’a jayga’-una ki bo moc a’-an} \]
\[ \text{REL place-LOC COMP many people be(PRS/FUT)-3PL I like(PRS/FUT)-ST-1SG} \]
\[ \text{‘I like a place where there are lots of people.’} \]

As Bashir (ibid.) observes, the last sentence has no pronominal correlate and formally resembles a free relative. However, as shown further in the dissertation (Bashir 1988, 367ff.), free relatives in Kalasha have a totally different structure; hence, such examples are best interpreted as having a null correlate.

Interrogative-indefinite affinity There is no data available for Kalasha interrogatives in conditionals. However, it is clear that the language at least does not have robust interrogative-indefinite affinity, and seems to license the indefinite reading of interrogatives in certain contexts only, e.g. when accompanied by the adjective ‘other’:

(28) \[ \text{war’ek k’aas putr se} \]
\[ \text{another who:OBL son that:SG:NOM} \]
\[ \text{‘That one (was) a son of someone else.’} \] (Di Carlo 2010, 80)
1.6.4 Shina [WH]

Readings One more rare example of an Indo-Aryan language with wh-correlatives is Shina, another Dardic language. In this language, correlatives can have both definite and universal readings:

(29) \([kesi \textit{myei madat thaw}] zu \textit{kone gaw}\).
\hspace{1cm} \text{REL.ANIM.ERG my help did that where went} \\
\hspace{1cm} 'Where did \textbf{the one who} helped me go?' \hspace{1cm} \text{(Hook 1997, 141)}

(30) \([thó-o jóo \textit{th-áa-e}] \textit{to asá b-úu}\).
\hspace{1cm} \text{YOU-AGP REL.INANIM DO-PRF-2.M.SG TOP that be-FUT.3.F.SG} \\
\hspace{1cm} '\textbf{Whatever} you did, \textbf{that} will be.' \hspace{1cm} \text{(Schmidt and Kohistani 2008, 92)}

Interrogative-indefinite affinity Judging from examples, Shina interrogatives can function as indefinites when embedded under conditionals:

(31) \([kudi \textit{jo lap ban jée-ji lah-áa-n-oe}] \textit{to}\).
\hspace{1cm} \text{where what morsel arrangement anyone-ABSUP find-IPFV-AUX-PRS-2.M.SG TOP} \\
\hspace{1cm} \text{mo-r at-á me-DAT bring-IMP} \\
\hspace{1cm} 'If you find an arrangement for \textbf{any} \textbf{food} from anyone \textbf{somewhere}, bring it to me.' \hspace{1cm} \text{(Schmidt and Kohistani 2008, 91)}

2 Uralic

2.1 Old and Modern Hungarian [WH]

Readings Modern Hungarian correlatives allow both universal and definite readings, as seen from the following examples from Lipták (2008):

(32) definite
\hspace{1cm} \text{[RC} \textit{Amit Mari tegnap főzött}, azt \textit{a levest nem ette meg}\).
\hspace{1cm} \text{REL.INANIM:ACC Mari yesterday cooked that:ACC the soup:ACC not ate PV} \\
\hspace{1cm} \text{János. John} \\
\hspace{1cm} 'John didn’t eat the \textbf{soup that} Mari cooked yesterday.' \hspace{1cm} \text{(Lipták 2008, 289)}

(33) universal
\hspace{1cm} \text{[RC} \textit{Akit szeret Mari}, azt \textit{meghívta a buliba}.\).
\hspace{1cm} \text{REL.INANIM:ACC loves Mari that:ACC invited the party-TO} \\
\hspace{1cm} 'Who(ever) Mari loves, she invited to the party.' \hspace{1cm} \text{(Lipták 2008, 293)}
Correlatives in Old Hungarian, however, differed from modern correlatives in two important respects. First, the pronouns used were not bare interrogatives, but had to be preceded by the particle *vala* (derived from the verb ‘to be’). This series of indefinites had various functions, including a free choice interpretation. Second, Old Hungarian correlatives could only have a universal interpretation:

(34) \[ vala-my zyleteny[k \, h\, ym \, nemzeth], \textit{azth koronkeed} \, \textit{wr} \, \textit{vala-rel.an}im be-born-fut.3\text{sg} \, \text{male issue that-acc age-dist \, lord} \, \text{ystenye}k \, \textit{aldozzad} \, \text{god-dat \, sacrifice-imp.2\text{sg}} \, \textit{\textquoteleft whatever male issue is born, that should always be sacrificed to God\textquoteright} \]  

(Bende-Farkas 2015, 13)

Bende-Farkas explicitly states that these structures yield to a conditional analysis of correlatives; all of them could be paraphrased as conditionals. We conclude that Old Hungarian correlative structures strongly resembled conditionals and had only universal readings. The transition to Modern Hungarian presumably involved a semantic extension to definite contexts combined with the replacement of *vala*-interrogatives by bare interrogatives.

**Interrogative-indefinite affinity** No data is available for Old Hungarian. Modern Hungarian is \([-\text{i}=\text{i}]\) according to Gärtner (2009, 13).

### 2.2 Udmurt [WH]

**Readings** Correlatives in Udmurt are based on the use of the wh-pronoun *kudiz* ‘which’ with an internal head, or the interrogatives ‘who’ and ‘what’. The examples in Belyaev (2012) from the Besermen dialect and the corpora demonstrate that both the universal and the definite readings are available:

(35) \[ kud-iz \, zor \, zor-i-z \, k\, \text{\textasciitilde k\, nuna}l], \, \textit{so} \, \textit{g\, \text{"a}lt-i-z \, \text{"s}r\, \text{"a}\text{"s}z\text{-ez}} \, \textit{\textquoteleft The rain that rained for two days washed out the road\textquoteright} \]  

(Belyaev 2012)

(36) \[ \textit{kin-len} \, \text{jonat-\text{"s}k-em-\text{-ez}} \, \textit{pot-e}, \, \textit{so} \, \textit{\text{"s}og-ges \, kat\, t\, ja-\text{"s}k-e \, heal-\text{-detR-\text{pr}}.3\text{sg}} \, \textit{\textquoteleft Who wants to be cured, (s/he) heals better\textquoteright} \]  

(Udmurt Corpus: \textit{Udmurt dunne}, 2007.10.19)

**Interrogative-indefinite affinity** Corpus examples show that Udmurt allows using interrogatives as indefinites in conditional sentences:
(37) *Ti kilidį, mar šu-emín vaškala-os-li*; «En vij; noš *kin* you have.heard what say-PRF ancient-PL-DAT NEG kill[IMP.2SG] and who
vi-iz ke, sud-e šuro-z».
kill-PST.3SG if judgement-ILL happen-PRS.3SG

‘Ye have heard that it was said of them of old time, Thou shalt not kill; and whosoever shall kill shall be in danger of the judgment.’ (Matthew 5:21)

lit. ‘You have heard that it was said to the ancients: Do not kill; and if *someone* (lit. *who*) kills, he will be subject to judgement.’

(Udmurt Corpus: *Matfeješ žeč ivor, 1990–2010*)

3 Turkic

3.1 Old Turkic [WH]

Readings Correlatives (and finite relative clauses in general) are not typical for modern Turkic languages, but a correlative construction did exist in Old Turkic, based on interrogatives:

(38) [*tavar kin ningu üklisä ] bäglik aŋar kärgäyür
wealth REL.Poss if.grows bāg-hood to.him befits

‘Whoever acquires much wealth, being a bāg befits him.’

(Erdal 2004, 500)

Erdal provides the following note on the universal use of Old Turkic correlatives: ‘In what appears to have been the primary use, the pronoun serves as a variable argument, the content of the main clause being understood to apply for any value of that variable’ (p. 499). While definite uses are encountered in texts, according to Erdal they are secondary to the universal ones, so this is in fact a borderline case. Note the use of a conditional verb form in the subordinate clause.

Interrogative-indefinite affinity Erdal (2004, 210) notes that ‘The set of [interrogative-indefinite pronouns] unites different functions (as happens with such pronouns in a wide variety of languages): They serve with interrogative content, given in the first line of the table, but also as indefinites, i.e. they can also signify ‘anyone’, ‘anything’, ‘anywhere’ and ‘any’ respectively.’ This description is ambiguous between [+i=i] and [%i=i]. However, judging from the translation of the pronouns with the any- series and the examples provided, all of which, when indefinite, seem to contain a conditional, modal, or negative operator, we tentatively conclude that Turkic had a non-robust interrogative-indefinite affinity.

4 Tungusic

4.1 Udihe [WH/DEM]

Readings Some Tungusic languages, among them Udihe (Nikolaeva and Tolskaya 2001), have both wh-correlatives and dem-correlatives. Most Tungusic languages are not described in sufficient detail to allow us to draw any conclusions on the semantics of correlatives. Udihe is exceptional in this respect due to the existence of two comprehensive studies: the grammar of Nikolaeva and Tolskaya (2001), which contains a large number of examples of correlatives but
does not discuss their semantics in enough detail, and Baek (2016), which includes an areal study of *wh*-correlatives in the languages of Eastern Eurasia (Eastern Siberia and Northern China), with a particularly detailed description of Udihe. While there is no explicit discussion of the semantics of Udihe *wh*-correlatives as such, all of Baek’s examples involve universal quantification. Furthermore, it is explicitly stated (p. 200) that the only two forms that can be used in subordinate correlative clauses are the simultaneous converb in -mi and the conditional converb in -(l)isi. According to the author, both involve conditional semantics (ibid.); the only difference is in switch reference. The following examples demonstrate *wh*-correlatives in Udihe:

\[(39) \ ilə olokto-mi, \ uta-la diga-i \]
\[\text{where} \ cook\text{-SIM.CVB} \ \text{that-Loc} \ eat\text{-PTCP.PRS} \]
\[\text{`He eats wherever he cooks.'} \quad \text{(Baek 2016, 201)}\]

\[(40) \ sii ono diana-isi-i, \ bii uta=bada nixə-i-mi \]
\[\text{thou} \ \text{how say-COND.CVB-2SG I} \ \text{that=PTCL do-PTCP.PRS-1SG} \]
\[\text{`I do however you say.'} \quad \text{(Baek 2016, 201)}\]

According to Baek, such *wh*-correlatives are not necessarily the result of Russian influence but rather a wider Eastern Eurasian areal feature (which might also explain their existence in Ket). Unfortunately, Baek does not discuss *dem*-correlatives. However, they are mentioned in the grammar of Nikolaeva and Tolskaya (2001). For our purposes, it is enough to show that, unlike *wh*-correlatives, they do have a definite reading, as exemplified by the following:

\[(41) \ dem\text{-based} \]
\[Nua-ni uti xoton-digi eme-gi-e-ni \ [# \ anana uta-du bagdi-e-i \:]. \]
\[he-3SG that city-ABL come-REP-PST-3SG you long.ago REL-DAT live-PST-2SG \]
\[\text{`He came from the same city you used to live in.'} \quad \text{(Nikolaeva and Tolskaya 2001, 666)}\]

Among the examples provided by Nikolaeva and Tolskaya, there is not a single one that involves a universal reading. Thus we can conclude with some certainty that the distribution of *dem*-correlatives and *wh*-correlatives corresponds to the distinction between definite and universal readings, as predicted by our hypothesis.

**Interrogative-indefinite affinity**  No data is available for Udihe.

## 5 Yeniseian: Ket [WH/DEM]

**Ket** (Yeniseian) has both *wh*-correlatives and *dem*-correlatives (Georg 2007; Nefedov 2012). While there is no explicit discussion of their semantics, the available examples show that *wh*-correlatives are used for universal RCs while *dem*-correlatives are used for definite RCs:

\[(42) \ dem\text{-based} \]
\[\text{[}qode kuŋa qaj bat d[il]-asan-l-bed \text{]} \ tu-de da-o-n-a-dij \]
\[\text{REL thou.DAT just PTCL 1-speak-PST-make this-F 3F-PST-PST-3COREF-reach} \]
\[\text{`That (woman) I was about to tell you about (just) showed up.'} \quad \text{(Nefedov 2012, 218)}\]
Note that the former construction is in fact indistinguishable from simple juxtaposition (‘I was about to tell you about that woman; she just showed up’), although the fact that it is discussed in reference descriptions such as Nefedov (2012) shows that it does have a degree of grammaticalization. The latter construction might, in fact, be a calque from Russian (Yulia Galyamina, p.c.), which makes this data point somewhat less reliable. Unfortunately, new data are difficult to obtain given that Ket is a moribund language.

**Interrogative-indefinite affinity**  No data is available for Ket.

### 6 Basque [WH]

**Readings**  Constructions that are described as correlatives in Northern Basque only have a universal reading (Rebuschi 2009). Rebuschi describes them as a subtype of conditionals.

(44) Bainan [nor-k ere irau-nen bai-du akhabantzaino], _eta hura_  
but REL_ANIM-ERG ever last-PROSP AUX end-until and DEM  
 _iza-nen da salbatua._  
  
`... but he that endureth to the end shall be saved.`  
(`Qui autem perseveraverit in finem, hic salvus erit.`)  
(Haraneder 1740, Rebuschi 2009, 94)

There are several features here that are noteworthy. First, the verb in the correlative clause is marked by the prefix _bait(t)-_, which, as Rebuschi argues, is the same element that marks conditional clauses, and also certain series of _wh_ -based indefinite pronouns. Second, the interrogative in the correlative clause must be accompanied by the particle _ere_, which is glossed as ‘ever’, but whose meanings include ‘also’ and ‘even’. Thus we are not dealing with a _bare_ _wh_-word in the subordinate clause; yet, as far as can be seen from the description, _ere_ in its use with interrogatives seems to be a specialized correlative marker. Third, the two clauses may also be connected by the conjunction _eta_ ‘and’:

(45) [Nor ere maite bait-dut ], _eta hura_ begira-tuko dut.  
  
REL_ANIM ever love _bait-AGT.3/EXP.1_ and DEM preserve-PROSP AGT.3/EXP.1  
  
`Whoever I love, them I will look after.`  
(Rebuschi 2009, 83)

This suggests a paratactic origin for this correlative structure. Among other things, this rules out an origin from left dislocation of a free relative or any other similar construction.

**Interrogative-indefinite affinity**  Basque is classified as [−i=i] in Gärtner (2009, 13), which formally classifies the language as an exception to our generalization. However, Basque correlatives are not ‘true’ _wh_-correlatives in the sense that they use an explicit additional marker _ere_. Hence, a special context for licensing interrogative-indefinites is not really required.
7 Dravidian

Dravidian languages employ two major strategies of relativization: participial RCs and correlative RCs. The relationships between these two types has been much discussed in the literature. The traditional view of Dravidian syntax is that each sentence must have only one finite verb, cf. e.g. Caldwell (1913, 488). In this context, correlative, a kind of finite subordination, were naturally viewed as a contact-induced feature that is due to Indo-Aryan influence. However, unlike Indic relative clauses, which use the *yo-pronominal series that originally probably had the demonstrative function (Frisk 1954–72, Chantraine 1980, s.v. hos), Dravidian languages use interrogatives as relativizing elements. Therefore, even if one assumes a degree of external influence, by and large this must have been an internal structural development. Therefore, Dravidian data are worth including in our typology alongside Indo-Aryan data. In this section, we will describe the relevant parameters for Old and Modern Tamil, two languages in this family for which correlatives seem to be described in most detail.

7.1 Modern Tamil [WH]

Readings  Modern Tamil has two kinds of correlative structures. In both of them the relative phrase is expressed by an interrogative and resumed by a demonstrative in the main clause. In one construction, the subordinate clause is finite and marked by the sentence-final particle =oo (Lehmann 1993, §§2.66, 4.86; Ramasamy 1981); see (46) for an illustration. Such constructions are always interpreted as definite (Thomas Lehmann, p.c.).

(46) [evan neerru va-nt-aan=oo] avan en tampi
rel.m yesterday come-pst-3sg.m=comp he my brother
‘He who came yesterday was my brother.’ (Lehmann 1993, ex. 358)

In the second construction, the verb in the subordinate clause stands in the non-finite conditional form, which is marked by the additive particle =um. In the absence of interrogatives, such clauses are interpreted as concessive (Lehmann 1993, §2.55). When an interrogative is placed in the subordinate clause and resumed by a demonstrative in the main clause, this construction becomes a typical universal correlative structure:

(47) [enta paiyan mutalil va-nt-aal=um], antap paiyanukku anta peenaavaik
rel boy first come-pst-cond=add that boy-dat that pen-acc
to
kotu give
‘Give the pen to whichever boy comes first.’ (Thomas Lehmann, p.c.)

Such clauses not only have universal semantics but are in fact syntactically indistinguishable from ordinary concessives with an interrogative-indefinite pronoun in the subordinate clause that is anaphorically connected to a main clause NP (see ex. 48).

As for definite correlatives, the data of Old Tamil (section 7.2) suggest that they only developed secondarily from conditional-based, universal correlatives.

---

9. Lakshmi Bai (1985) appears to claim that all correlatives in Tamil have the universal/indefinite reading. This presumably semantically differentiates them from participial relatives, which can have both definite and universal readings. However, other data on Modern Tamil do not confirm Lakshmi Bai’s generalizations (Thomas Lehmann, p.c.).
Interrogative-indefinite affinity Interrogatives in Tamil can be interpreted as indefinites in concessive clauses, even in the absence of any overt or covert correlate:

(48) \textit{yaar kuuppitt-aal=um, naan pook-a maatt-een}  
who call-COND=ADD I go-INF will  
'No matter \textit{who} calls, I won’t go.' 
(Thomas Lehmann, p.c.)

7.2 Old Tamil [WH]

Unlike Modern Tamil, correlatives were very rarely used in Old Tamil. The overwhelming majority of examples involve concessive clauses and have a universal interpretation:

(49) \textit{[etticai cel-in=um ] atticai cpru}  
REL.direction go-COND=ADD that.direction rice  
'\textit{Whatever direction} we go that direction has rice.'  
(Puṟanaṇṟu 206.13, Thomas Lehmann, p.c.)

Like in the modern language, these clauses cannot be clearly separated from ordinary concessives or conditionals with \textit{wh}-indefinites.

Correlatives that did not utilize conditional forms are very rarely encountered in texts. One such instance is with nonverbal predicates (Ramasamy [1981]), which is explained as the consequence of nominals lacking non-finite forms (or any forms with verbal morphology, for that matter) in Lehmann (1998).

Therefore, we can conclude that in Old Tamil conditional correlatives were overwhelmingly predominant, and other types of correlatives were yet in a nascent state. It is difficult to draw definitive conclusions without a comprehensive study of both Old and Middle Tamil, but one can hypothesize that the extension from conditional to definite readings occurred gradually during later stages of the languages via a shift to other forms in dependent clauses.

8 East Caucasian

In East Caucasian, the main relativization strategy is invariably the participial prenominal gap construction. However, correlatives do exist in some languages, and for the most part, they have only a universal reading.

8.1 Kryz

Readings Authier (2009, 377) notes that in Kryz ‘[d]ans les relatives à pronom, les modes hypothétiques sont les plus fréquent, servant à former les relatives ‘indefinies’. He calls this type of correlatives \textit{relatives conditionnelles} and provides examples like the following:

(50) \textit{[hal-ir va-z kar taklif ar-na ], hala adami.c-zina y-i-ghin!}  
REL.ANIM-ERG 2-DAT work invitation do-COND that person-INS pv-draw.IMP  
'Si quelqu’un te propose du travail, va avec lui.' 
(Authier 2009, 378)

Observe the use of the conditional in the above example: according to Authier, it is obligatory in universal correlatives. He notes that the definite reading is also available, when indicative
mood is used. However, in practice, all of his examples represent contexts that are ambiguous between a definite and a universal reading, as they involve unique referents. Consider the following example:

(51) \textit{hal-id q’ay-cu a-d an q’ay-i!} \\
\hspace{1cm} \textit{REL.Poss-HUM die-PREF.F 3-HUM AN die-OPT} \\
\hspace{1cm} ‘Celui de qui elle est morte, qu’il meure aussi!’ \hspace{1cm} (Authier 2009, 380)

In this example, the possessor of the implicit subject of the verb ‘die’ is obviously unique, hence, the quantificational interpretation and the definite interpretation are extensionally equivalent. Therefore, at best, Kryz has both semantic types of correlatives differentiated by mood morphology. But in fact, it is more probable that the universal/conditional reading is the only one available.

**Interrogative-indefinite affinity** Kryz, in addition to correlatives, has bare interrogatives in conditional clauses, as in the following example:

(52) \textit{[hal-ir riki ğe-t’-a-na], nabada riki açuğ yi-yu vun!} \\
\hspace{1cm} \textit{who-ERG door PV-strike-COND God.forbid door open do-DEONT 2} \\
\hspace{1cm} ‘Si qui que ce soit frappe à la porte, garde-toi bien de (lui) ouvrir!’ \hspace{1cm} (Authier 2009, 380)

In this example, the correlate is not overtly expressed, and the sentence can just as well be interpreted as a conditional: ‘If someone knocks at the door, God forbid you to open!’

### 8.2 Khwarshi [WH]

**Readings** The syntax of correlatives in Khwarshi (Tsezic) is similar. According to Khalilova (2009, 357), such correlatives use interrogatives, and the verb in the subordinate clause stands in the General tense, which has habitual semantics:

(53) \textit{[q’arλ’a hibo b-āh], izzu-qo golluč himon} \\
\hspace{1cm} \textit{early REL.ANIM HPL-stand.GNT that.PL.(P)OBL-CONT everything thing(G4).ABS} \\
\hspace{1cm} l-ēqʷ G4-happen.GNT \\
\hspace{1cm} ‘One who gets up early manages to do everything.’ \hspace{1cm} (Khalilova 2009, 357)

This implies that correlatives are universal, although this is not explicitly stated in the grammar.

**Interrogative-indefinite affinity** Khalilova (2009) does not discuss the subject explicitly, but notes that the interrogative \textit{doco} ‘(how) many’ can be used not only in interrogative clauses, but also as an indefinite in concessives:

(54) \textit{[doco O-ul’-še O-eč-lon], uže isuyol O-ôk’-i} \\
\hspace{1cm} \textit{many G1-be.afraid-IPFV.CVB G1-be-CONC boy(G1) that.APUD.LAT G1-go-PST.W} \\
\hspace{1cm} ‘No matter how (he) was afraid of him, the boy went to his place.’ \hspace{1cm} (Khalilova 2009, 154)

This suggests that Khwarshi is [%i=i].

19
8.3 Tsez [WH]

Readings  Interrogative-based correlatives are described for Tsez (Tsezic) in Polinsky (2015, 293–298). No discussion of their semantics is provided. However, all examples permit quantificational interpretation, and many of them are explicitly translated with whoever type pronouns. Notably, there seems to be no strict requirement for a correlate in the main clause, implicit or explicit. According to Polinsky, the following sentence ‘is ambiguous in referring to the dog and its owner’ (p. 298):

(55) \[
\begin{array}{llllllll}
\text{nāzo} & \text{uži-z} & y^\text{way-ā} & \text{mi} & \text{han-ā} & \text{šul-} & \text{ir} & \text{ža!} \\
\text{REL.OBL} & \text{boy-GEN2} & \text{dog-ERG} & \text{2SG.ABS} & \text{bite-PST.W.INTERR} & \text{FORGET-CAUS.IMP} & \text{DEM.ABS} & \\
\end{array}
\]

‘Whichever boy’s dog bit you, forget it/him.’ (Polinsky 2015, 298)

Interrogative-indefinite affinity  Polinsky (2015) does not provide any specific examples, but her discussion of interrogatives suggests that Tsez is [%i=i]: ‘Interrogative words are part of the set of indeterminate (indefinite) lexical items, which means that they can have an interrogative proper interpretation or an indefinite interpretation (for further discussion of the indefinite series, see Ch. YY[Particles]). For example, the word šebi ‘what/who’ can also mean ‘something’, and its interpretation depends on the nature of the clause in which it occurs.’ (Polinsky 2015, 185)

8.4 Standard Dargwa [WH]

Readings  Mutalov (2017) describes two correlative constructions in Standard Dargwa, both wh-based. In one, which he calls general correlative (obščekorreljativnaja), the verb in the subordinate clause has the conditional / concessive suffix -alra; the construction has a universal interpretation:

(56)  general correlative

\[
\begin{array}{lll}
\text{quli} & \text{či} & \text{uhnawq-[alra]} \\
\text{room.LOC[LAT]} & \text{REL.ANIM} & \text{M:COME.IPFV-CONC.3SG} \\
\text{herik'uli} & \text{sa-j} \\
\text{M:LOOK.IPFV:CVB} & \text{COP-M} \\
\end{array}
\]

‘Whoever comes to the room, he looks at the picture.’ (Mutalov 2017, 73)

The second construction, which Mutalov calls specific correlative (častnokorreljativnaja), has a similar form, but the subordinate clause carries the indirect question or uncertainty suffix -il (and morphophonological variants thereof). The origin of the indirect question marker and its original function, as well as its relationship to other similar suffixes including the conditional suffix, is unclear.

(57)  specific correlative

\[
\begin{array}{lllll}
\text{darsli-či} & \text{či} & \text{q'aniʔubl-[il]} & \text{il} & \text{ajzab} \\
\text{lesson-SUPER[LAT]} & \text{REL.ANIM} & \text{BE.LATE.PVF:CVB-1Q} & \text{M:STAND.UP:OPT}[3] \\
\end{array}
\]

‘Who is late to the lesson, let him stand up.’ (Mutalov 2017, 73)
According to Mutalov, the difference between the two constructions is that general correlatives always have a universal reading, while specific correlatives refer to definite individuals. Unfortunately, all examples provided for specific correlatives have an ambiguous interpretation, so in this case we can only rely on the author’s intuition as a native speaker. It is possible, however, that specific correlatives actually allow both readings, but gives the impression of being definite-only through opposition to the universal construction.¹⁰

Interrogative-indefinite affinity  In the same article, Mutalov (2017) demonstrates that so-called general correlatives are formally no different from normal conditional clauses. In particular, they can function without a correlate, as in the following example:

(58) [či wak’-alra ] iš ra ziʔirar
who m:come.1PFV:CONC.3 this be.happy.1PFV:hab.3
‘Whoever, comes, he j/*i (this one) is happy.’ (Mutalov 2017, 73)

The pronoun iš used in the main clause in (58) is a proximal demonstrative that can only have a deictic meaning and cannot refer to a referent mentioned earlier in the discourse. Therefore, it refers to a different person that the one who enters the room. This example shows that conditional clauses license the indefinite reading of interrogatives in Standard Dargwa.

9  Sino-Tibetan

9.1  Chinese [WH]

Readings  It is well-known that Chinese possesses a correlative-like construction where both NPrel and NPMat are interrogative:

(59) [shuí xiān lái ], shuí xiān chī
who first come who first eat
‘Whoever first comes first eats.’ (Luo and Crain 2011, 756)

It is not clear whether such constructions belong to the class of correlatives proper. In any event, they are outside of the scope of this paper because they clearly result from a different diachronic scenario, unless one assumes that interrogatives replaced demonstratives/personal pronouns at some stage (for which there is no evidence). However, Chinese does also allow a more conventional kind of correlative structure:

(60) [kàn shuí kēwù], jiù duŏkāi tā
see who hateful then avoid him
‘I avoid whomever is hateful.’ (Lu 2002, 94, via Baek 2016, 198)

Two features of this structure are important for our analysis. First, these correlatives only have a universal reading; in fact, Cheng and Huang (1996, 150) analyze them as a subtype of conditionals with ruguo ‘if’ omitted. This is related to the second important observation, namely, that it is obligatory to use jiù ‘then’ in the main clause in such correlative sentences. This

¹⁰. Tamil correlatives might be treated as universal-only in Lakshmi Bai (1985) for a similar reason, viz. due to their opposition to definite-only participial relative clauses; in fact, correlatives in Tamil can have both readings. For more information, see section [7]
means that these correlatives are either historically derived from or synchronically identical to conditional sentences; the interrogative serves as an indefinite element.

**Interrogative-indefinite affinity**  Chinese is listed in Gärtner (2009, 13) as a [%i=i] language.

### 9.2 Tibetan [WH]

**Readings**  Tibetan correlatives are based on interrogatives in the subordinate clause and necessarily involve the marker *na* ‘if’ (Cable 2009). Cable shows that Tibetan correlatives can have both the definite and the universal readings. The sentence in (61), for instance, is translated in the universal sense; yet Cable also notes that this is only a convention that he follows, and that the sentence ‘is reported as felicitous in contexts in which there is a single yak which the addressee has purchased, and the speaker wishes to communicate that they have killed that particular yak’ (Cable 2009, 199).

(61)  
\[
\text{Khyodra-s gyag gare nyos yod na} \ ng\text{-s de} \ bsad pa \ yin.
\]

you-\textit{erg} yak what buy AUX if I-\textit{erg} that kill PRF AUX

‘I killed whatever yak you bought.’

‘I killed the (particular) yak you bought.’  (Cable 2009, 198)

**Interrogative-indefinite affinity**  There is no conclusive data on interrogative-indefinites in Tibetan. However, since Tibetan correlatives use an explicit conditional element *na* ‘if’, the link to conditional-licensed interrogatives is transparent. Also worth noting is that Gärtner lists Newari, another Tibeto-Burman language, as a [%i=i] language.

### 10 Niger-Congo

Correlatives are a dominant strategy in several West African languages belonging to the Mande and Gur language families. In principle, correlatives in these languages have both definite and universal readings, but the way the universal reading is expressed supports our hypothesis.

#### 10.1 Mande [DEM]

**10.1.1 Both readings available**

In Kita (Mande > Niger-Congo), the correlative marker *mín* is derived from a demonstrative, according to Creissels (2009). The universal reading is possible, but only in the presence of the ‘if’ marker, otherwise absent:

(62)  
\[
\text{[ídi wori mín tà ] à dú n mà!}
\]

2ACP.POS money.DEF REL take 3 give 1SG POSTP

‘Return to me the money that you’ve taken!’  (Creissels 2009, ex. 14d)
The same situation obtains in Mandinka (Mande), where the correlative is marked by cognate mëŋ. Unmarked correlatives only have a definite reading, and the subordinator ‘if’ must be used to get the universal reading:

(64) definite

[Mus-ôo mëŋ ña nins-ôo jé ], wo tâa-tâ a ŋinëŋ
woman-D REL ACP.POS COW-DEF see DEM go-ACP.POS 3SG search
‘The woman who had seen the cow went looking for it.’

(Creissels and Sambou 2013, 465)

(65) universal

[Nîŋ mëŋ kó a lafi-ta ñ jé-la ], a ŋinëŋkäa fóló
if person.DEF REL QUOT 3SG want-ACP.POS 1SG see-INF 3SG ask first
a tóo la
3SG name.DEF OBL
‘Whoever says s/he wants to see me, ask him his name first.’

(Creissels and Sambou 2013, 475)

Thus we find that in all these languages, basic correlatives only have a definite reading. Universal ‘correlatives’ are effectively a subtype of conditionals, and their grammaticalization represents a subtype of the scenario that we are proposing in the paper. Namely, a full, non-pronominal NP is interpreted as an indefinite within the scope of a conditional, and resumed by a demonstrative or another referring expression in the main clause. The only difference from wh-correlatives is that in the latter, only the indefinite reading is available and only in the scope of a conditional, making the universal reading primary. No such constraint applies to noun phrases, therefore, in such languages both definite and indefinite readings are possible; they are only distinguished by the marking of the subordinate clause.

10.1.2 Definite reading only

While most Mande languages of West Africa allow both definite and universal readings for correlatives, but the latter require additional conditional marking. But in some languages of this family, the universal reading seems to be altogether unavailable for correlatives, and it is expressed using entirely different constructions. A case in point is Bambara:

(66) definite

[ nèğe mín dïra à mà ], ô dë yë nïn yë
‘The ring which was given to him, it is this one.’

(Dumestre 2003, 367)

11. Dumestre (2003) and Welmers (1976) do not provide any gloss, and we have decided to refrain from providing our own due to insufficient knowledge of Bambara. However, the identification of the relevant components of the correlative construction (highlighted in bold) is secure enough.
universal

\[
\text{màrifa bè } mògò \text{ ó } mògò \text{ dè bólo }, \text{i bè màrifa cì i fāso dè kɔsɔn}
\]

'Every person who has a gun uses it for their homeland.' (Dumestre 2003, 369)

According to Dumestre’s description, definite restrictive relative clauses are usually encoded by correlatives (66). Relative clauses with the semantics of universal quantification, however, utilize a distinct construction which the author calls ‘proposition impliquante à relateur intra-propositionnel’ (implicative propositions with the inter-propositional connective \(\text{ó} \)). The syntax of these latter is completely different from that of correlatives: they use a different marker \(\text{ó} \), which connects two tokens of the same noun (e.g. \(mògò \text{ ó } mògò \) ‘every person’ in 67 above). From this description and the examples provided one can conclude that correlatives in Bambara cannot have the universal reading.

According to Diagana (1995, 465–466), correlative relative clauses in Soninke (Mande), whose marker is also derived from a demonstrative, are always definite, both formally and semantically:

\[
\text{ó gà dà gōlē } \text{ ké } \text{ bé } \text{ dəbərə lə’k}i \text{ ] } \text{ á } \text{xōtē } \text{n } \text{ yà } \text{n} \]

The work that we have done today is difficult.’ (Diagana 1995, 463)

Finally, while it is not stated explicitly, the description of Vai (Welmers 1976) strongly indicates that correlatives in this language are only used in definite contexts:

\[
\text{kàì mú } \text{’à à mà } \text{], } \text{à bè n}iē
\]

The man who did it is here.’ (lit. ‘Man top did it, he is here.’) (Welmers 1976, 124)

10.2 Senufo: Supyire [DEM]

Correlatives in Supyire (Senufo) are almost identical to those in Kita and Mandinka. Here, as well, unmarked correlatives have only a definite reading, and universal sentences must use the conditional marker \(kā\):

12. Carlson (1994, 488) links the correlative marker \(kè\) to the identical locative question marker \(kè\) ‘where’. However, he also notes that ‘[b]oth the relative \(kè\) and interrogative \(kè\) are quite likely derived from the gender 2 singular demonstrative \(b̥kè\)’. Considering the data of other languages of the region with similar constructions, including Bambara, a direct diachronic connection between the relative and demonstrative marker seems more likely than a development from interrogative ‘where’ to relative.

13. Carlson actually distinguishes between ‘conditional’ and ‘distributive’ relativization; the latter, he claims, may be expressed by the normal correlative construction. However, all of the examples provided can be construed as being definite. Truly universal relative structures seem to require conditional marking. A good case in point is 71, where the interpretations ‘Whatever child she got would die’ and ‘Those children she got used to die’ have rather different implications.
definite

\[
[\text{Yaagé } \text{ka } \text{a ū bò ké }], \mu \text{a kùrù cé.}
\]  

thing.def it PRF him kill REL you PRF it(emph) know

‘You know the thing that killed him.’  

(Carlson 1994, 488)

universal

\[
[U \text{ahá pyáŋi ſgé-mù tà ké }], \text{ura asi ṅ-tòrò.}
\]  

she cond child.def dem-rel get REL it(emph) hab.seq intr-pass

‘Whatever child she got would die.’  

(Carlson 1994, 508)

Thus, just as in the Mande case described above, basic correlative clauses in Supyire only have the definite reading, and the universal correlative clauses are a subtype of conditional clauses.

References

Arsenijević, Boban. 2009. “[Relative {conditional {correlative clauses}}].” In Lipták 2009, 131–156.


