

A Proposal for a More Universal Annotation of Relative Clauses in Universal Dependencies

Santiago Herrera¹, Sylvain Kahane^{2,3}

¹Université Sorbonne Paris Nord, LIPN, CNRS, UMR 7030

²Université Paris Nanterre, Modyco, CNRS ³IUF - Institut Universitaire de France
herrera@lipn.univ-paris13.fr, skahane@parisnanterre.fr

Abstract

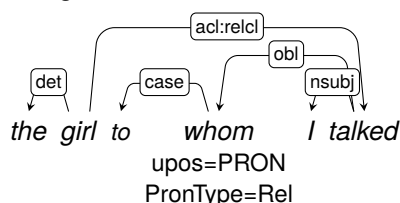
This paper proposes a new encoding of relative clauses compatible with the Universal Dependencies annotation scheme for syntactic treebanks. After showing that the current guidelines are based on the main strategy of relativization in European languages, we show that it cannot be easily extended to other relativization strategies, especially head-internal relatives clauses. The criteria for the POS of relativizers are discussed. We apply our annotation to an important variety of strategies in Mandarin, Japanese, Bambara, German, Basque, Turkish, Latin, Beja, Gbaya, and Wolof, as well as to various strategies in English, including participial clauses.

Keywords: relativization strategy, relativizer, head-internal relative clause, head-external relative clause, free relative clause, participial clause, language typology

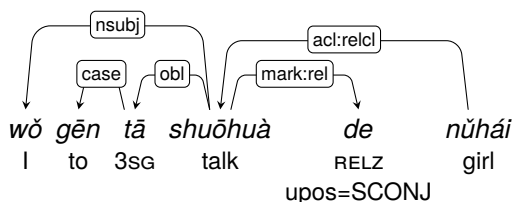
1. Introduction

The current UD annotation of relative clauses (henceforth RCs) is primarily based on the relation `acl:relcl` between the RC and its antecedent, and on the `PronType=Rel` feature on the relativizer, when it is a relative pronoun (`PRON` in (1a) for English).

(1) a. English



b. Mandarin



‘the girl I talked to’, lit. the girl that I talked to her

This annotation has been mainly thought for languages where the RC has a relative pronoun (`PronType=Rel`) and depends on a nominal antecedent (`acl:relcl`).¹ In some cases, such as

Authors are listed in alphabetical order.

¹This is confirmed by the definition of RCs given on the UD website: “A prototypical relative clause (RC) modifies a nominal that is understood to fulfill some grammatical role in the RC. The head is said to be ‘extracted’ from the RC.” (Universal Dependencies, 2026d)

in (1b) for Mandarin, the relativizer (`SCONJ`) is marked by the relation `mark:rel`. This usage is not widespread across treebanks, which usually annotate relativizers with `PronType=Rel` feature. However, when the RC is marked by a pure relativizer, the currently proposed encoding is not very satisfactory either. The `PronType=Rel` feature on a word analyzed as a subordinating conjunction can be misleading, and the relativized position is not indicated, whether it is a gap or it is occupied by a resumptive pronoun, such as *tā* in (1b).

Finally, encoding the fact that the clause is a RC on the label of the relation between the clause and its antecedent (`acl:relcl`) is also debatable, as the property of being a RC is a property of the clause itself rather than its relation with its antecedent.

Relative clauses with an antecedent and a relative pronoun are far from the only strategy for relativization (Keenan and Comrie, 1977; Downing, 1978; Lehmann, 1986; Comrie, 1998; Croft, 2022). It is, in fact, at one end of the continuum between all the possibilities for building a relative clause, even if it is a common one in European languages.

For this reason, and in order to be able to compare the different RC structures across languages, typologists work with functional or semantic definitions of the relativization phenomenon. From a functional perspective, a relative construction is an action that modifies a referent (Croft, 2022). Prototypically, a RC specifies or restricts a set of objects (Keenan and Comrie, 1977) denoted by a noun phrase in the matrix clause. This definition narrows RCs down to those that restrict the set of objects being referred to, while allowing for better cross-linguistic comparison.

In this paper, we propose a review of RC strate-

gies and a new encoding that can be adapted to the various strategies of relativization. We start our review at the other end of the spectrum, by the head-internal strategy.

Section 2 reviews recent contributions within the Universal Dependencies framework aimed at improving cross-linguistic comparisons of constructions. Additionally, it examines the current state of RC annotation in its last release (v2.17, Zeman et al., 2025). In section 3, we present the head-internal RCs in Japanese and Bambara, as well as the free RC of English. In section 4, we revisit the familiar strategy of head-external RCs, exemplified by English and, for a less common variant, by Gbaya. Sections 5 and 6 focus more on the relativizer itself. We start by the case of relativizers that are part of the verbal form, such as in Basque and Turkish, as well as some constructions of English. Section 6 is devoted to the analysis of relativizers which are separate forms and especially to the choice between `PRON` and `SCONJ`; it is illustrated by Latin, Beja, Wolof, and again English.

Our contributions can be summarized in three points:

- we provide an analysis of existing RC annotation, identifying a bifurcation between UD guidelines and actual annotation practices;
- we propose cross-linguistic analyzes for a more consistent annotation for different RC strategies;
- we introduce new tags that ensures comparative annotation without information loss.

2. Research Context

The UD collection has been growing in both size and diversity, and has become a priceless resource for quantitative syntactic typology (e.g., Levshina, 2019; Gerdes et al., 2021; Levshina, 2022) and other larger cross-linguistic studies, even for those typological studies that do not work directly with UD data, but use parsers that have been trained on it (e.g., Östling, 2015; Kann, 2024). In this context, comparability and typological reliability become crucial.

2.1. Related work

Continuous efforts are being made in UD to improve coverage of language diversity within a comparable framework, independent of language-specific strategies. One research line has focused on evaluating and revising UD annotation through the concept of comparative constructions (Croft et al., 2017; Nivre, 2025). Recently, these authors have started a review of the compatibility between the UD framework and linguistic typology

principles, in order to identify limitations and propose possible solutions. To date, the authors have studied reference and modification constructions (Nivre and Croft, 2025) and general verbal predication constructions (Croft and Nivre, 2025). Other work has focused on annotating constructions atop UD (Weissweiler et al., 2024) that could then be compared cross-linguistically. As far as we know, RCs have only been annotated as constructions in the EWT English Treebank (Silveira et al., 2014); the role of the relativized element is indicated even in cases where there is a gap, based on enhanced dependencies. Our work builds on the aforementioned authors' work, offering an immediate practical perspective. We revisited UD annotation, focusing only on relative clauses. Without proposing a different layer of annotation, we propose new features that are necessary to provide the information required for cross-linguistic comparisons of RCs. We argue that, given the current annotation, such a comparison is unfeasible. In addition, this work serves as a prerequisite for annotating RC constructions in the style of Weissweiler et al. (2024).

2.2. Relative Clauses in UD treebanks

Version 2.17 of UD comprises 339 treebanks for 186 languages. Analysis reveals that 21% of treebanks (with $\geq 1k$ tokens = 210 treebanks) do not contain an explicit RC relation (`relcl`) or a relative pronoun feature (`PronType=Rel`) within adnominal clauses (`acl`). This represents approximately 18% of the largest languages in UD, including Basque, Japanese, and Turkish. Even where annotation exists, retrieval of RCs is not straightforward. For instance, the Serbian and Croatian treebanks use the tag `PronType=Rel, Int`, combining relative and interrogative pronouns and making explicit filtering impossible.

While this absence could be attributed to inconsistent annotation practices, perhaps due to the optional status of relation subtypes, it suggests a general shortcoming across languages. The current preference in UD guidelines for strategies that use relative pronouns creates a representational imbalance. For example, languages using mainly participial or nominalized clauses (e.g., Turkish or Basque) lack the appropriate tags to annotate these constructions.

Consequently, annotators are driven to find workarounds for the scheme. We observe the feature `PronType=Rel` being forced onto non-pronominal categories like `SCONJ`, `AUX`, and `PART` in 15% of languages, simply to preserve the relative interpretation,² or the use of the more generic

²The guidelines indicate that `PronType=Rel` should be used for relative pronouns, determiners, numerals, and adverbs. (Universal Dependencies, 2026c)

`acl` label that obscure the specific syntactic function.

Ultimately, the current scheme fails to fully account for internally headed relatives, free relatives, or gapping strategies where the head is structurally integrated into the subordinate clause. For 40% of languages, it is not possible to identify the relativized element as they do not overtly indicate the relativizer (specifically when this is not an `SCONJ` or a `PART`). This results in an asymmetry of information encoding: only languages with explicit relative words (`PronType=Rel`) can fully encode the role of the relativized element, while other strategies are structurally misrepresented.

3. Head-Internal Relative Clauses

The Head-Internal RC (HIRC) strategy was first described by Kuroda (1974) for Japanese, and confirmed for Diegueño (now called Kumeyaay, Yuman) by Gorbet (1974) and for Navajo (Na-Dené) by Platero (1973). In HIRCs, the relativized noun stays inside the RC. Examples of HIRCs are now documented across a diverse range of language families. In this section, we first illustrate this phenomenon with Japanese and then move to Bambara (Niger-Congo).

Japanese complement clauses are built by adding the nominalizer affix *-no* to a finite clause.³ This is illustrated by (2b) where *-no* is added to the clause in (2a). This nominalized clause occupies the object position and receives the accusative marker *-o*.

(2) Japanese (Creissels, 2006)

- a. Ano hito-ga Nihon-e itta
DEM man-NOM Japan-to went
'This man went to Japan'
- b. [Ano hito-ga Nihon-e itta]-no-o
DEM man-NOM Japan-to went-NMLZ-ACC
shirimasen deshita
didn't know
'I didn't know that this man went to Japan'

The same strategy can be mobilized for RCs. In (3a), the verbal form *tukameata* takes *neko-o* as its object. The noun *neko* in (3a) can be replaced by the whole clause (3b) nominalized by *-no*, giving (3c). As the verb *tukameata* needs an entity as an object, the nominalized clause is now interpreted as an HIRC. The relativized position is not explicit

³The segmentation into words for Japanese is widely debated. We follow here the segmentation proposed by most typologists, which more or less corresponds to the traditional segmentation into *buntetsu*.

and the HIRC is potentially ambiguous, both arguments of *tukameata* being able to be the target.

(3) Japanese (Creissels, 2006)

- a. Neko-o Hanako-ga tukamaeta.
cat-ACC Hanako-NOM caught
'Hanako caught a cat.'
- b. Neko-ga sakana-o torootoshita.
cat-NOM fish-ACC tried_to_rob
'A cat tried to rob the fish.'
- c. [Neko-ga sakana-o torootoshita]-no-o
Hanako-ga tukamaeta.
'Hanako caught a cat that tried to rob the fish.'

It is inconvenient to annotate that there is a RC in (3c) given the current UD annotation of RCs. First, the relativizer is generally analyzed as a verbal affix and the feature `PronType=Rel` can hardly be associated to a verbal form. Second, there is no antecedent,⁴ therefore the `relc1` should be added directly to the relation between the verb *tukameata* and its object.⁵ Third, the relativized position would not be indicated. Consequently, we propose a completely different annotation:

- The fact that the clause is interpreted as an RC will be indicated by a clause-level feature `ClauseType=Rel` on its head.
- The position of the relativizer is indicated by a feature `Relz=Affix`, which replaces the feature `PronType=Rel`. Other values of `Relz` we consider are `Word` and `Stem`.
- The relativized location can be indicated by a feature `RelPosition=Yes`.⁶

Our analysis for (3c) is shown in (4). Note that the relative clause is analyzed as an `obj` of the main verb, because it refers to an entity. It contrasts

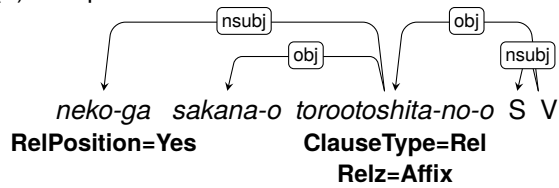
⁴Some authors, such as Murasugi (1991), analyzes *no* as a pronoun antecedent of the clause. In any case, we present other examples where the relativizer is a verbal affix in section 5.

⁵As we said in section 1, to be a RC is more a property of the clause itself than of its relation with its governor, although, in this case, it is the semantic properties of the verb governing the RC that triggers the RC interpretation. Encoding this information on a relation would involve adding `relc1` to a great number of relations. For other advantages of using a clause-level feature, see section 5 and 7.

⁶Here again, as suggested by one of our reviewers, it would be possible to add an subtype extension on the relation to indicate the relativized position. But it is really a property of the position itself and not of the relation with its governor.

with the subordinated clause in (2b) that refers to a proposition and would be `ccomp`.

(4) Japanese



Our second example is Bambara. UD has a small treebank of Bambara that contains 80 RCs (Aplonova and Tyers, 2017). Bambara has a unique relativizer, *min*, that can occupy any nominal relativized position or can be associated with a noun occupying the relativized position.⁷ The structure of the RC is identical to the structure of a main clause, the only difference being the additional relativizer *min* in situ. RCs in Bambara are almost always fronted. They are associated with an overt pronoun in the matrix clause if they occupy an argument position, as in (5a) and (5b), but not if they play an adverbial role, as in (5c). In examples (5a) and (5c), *min* is associated with the relativized noun (and can be analyzed as a determiner), while in (5b), it occupies the relativized position by itself.

(5) Bambara (v2.17, Aplonova and Tyers, 2017)

- a. a mana den **min** wolo, a
3SG COND.AFF child RELZ give_birth, 3SG
bè o dun
IPFV.AFF it eat

'He ate the child she gave birth to', lit. she gave birth to RELZ child, he ate it

- b. e ye **min** don nin ba
2SG.EMPH PFV.TR RELZ know this mother
ye o don.
PFV.TR it know

'This mother knew who you knew'

- c. dennin nana tuma **min**, a
child.DIM come.PFV.INTR moment RELZ, 3SG
ye foli kè
PFV.TR salutation do

'He saluted when the child arrived'

In the current annotation of Bambara treebank (v2.17, Aplonova and Tyers, 2017), the RC is analyzed as a dependent (`acl`) of the overt pronoun in the matrix clause. But the overt pronoun of the Bambara constructions is an ordinary pronoun that cannot be modified. It is equivalent to the resumptive pronoun in a dislocation (which is a common

⁷*min* can be pluralized, giving the form *minw*, which occurs in 6 of the 80 RCs.

construction in Bambara) and the construction is rather considered a correlative construction (Culy, 1990; Keenan and Comrie, 1977; Creissels, 2006; Croft et al., 2017). The link between the RC and the pronoun could be indicated by a coreference index (see Nivre and Croft (2025) regarding difficulties with coindexation in the UD). Here we propose to add a feature `RelAnt=Yes` on the pronoun in the matrix clause and to attach the RC to the main verb by a *dislocated* relation (Figure 1).⁸

In some sense, the HIRCs of Bambara are not noun modifiers, but rather free relatives coreferent or not with a pronoun. It must be noted that Japanese HIRCs are also close to free RCs; it is possible to suppress the noun in the relativized position, which becomes a gap, to obtain a meaning equivalent to a free RC in English.⁹

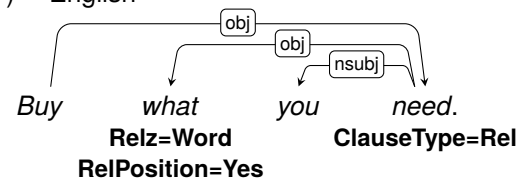
(6) Japanese

[Sakana-o torootoshita]-no-o
[fish-ACC tried_to_rob]-RELZ-ACC
Hanako-ga tukamaeta.
Hanako-NOM caught

'Hanako caught what tried to rob the fish.'

Many authors have made a link between HIRCs and free RCs and it is possible to analyze them in the same way we have analyzed HIRCs (Basilico, 1996; Grosu and Landman, 1998). In (7), *what* is a relative pronoun that fills a position inside the clause (as *min* in (5b)). As for (4), the relative clause, which refers to an entity, is `obj` of *buy*, even though we propose not to take *what* as an external head.

(7) English



This analysis is different from the current UD analysis, where the relative pronoun is annotated as the head, arguing that "Free relatives are nominals containing a relative clause where the relativizer is 'fused' with the head of the whole nominal" (Universal Dependencies, 2026d). Some authors (Lehmann, 1986; Le Goffic, 2002) defend a reverse analysis, considering that RCs with an antecedent have developed from the combination of

⁸We can also mention the relation `acl:crel` that has been proposed for UD_Sanskrit-Vedic treebank (Universal Dependencies, 2026a).

⁹Grosu and Hoshi (2016) analyze this construction as a "gapped light-headed HERC", considering that *no* is more pronominal in this case.

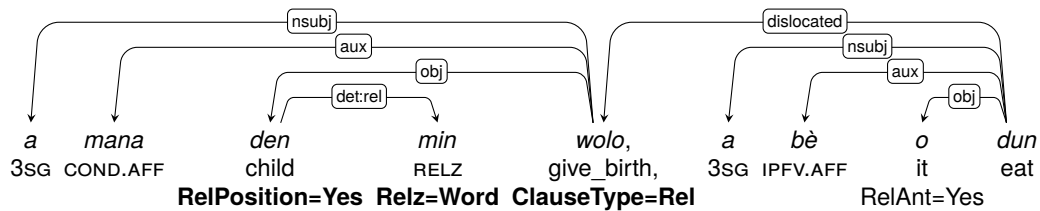


Figure 1: A correlative analysis for the Bambara RC.

a noun and a free relative occupying together the same position. For this reason, and to homogenize annotation, we believe it would be relevant to analyze free relatives in the same way as HIRCs.

We must nevertheless mention an additional problem, the case conflict (Bresnan and Grimshaw, 1978; Vogel, 2003), which is well illustrated by German. In German free RCs, the relative pronoun receives the case addressed to the relativized position, which proves it occupies this position inside the RC. But this case must also correspond to that assigned by the matrix governor to the entire RC: In (8a), the nominative relative pronoun *wer* is the subject of the RC and the RC is the subject of the main clause, while in (8b), the accusative relative pronoun *wen* is the object of the RC and the RC is the object of the main clause. But in (8d), the relative pronoun *wer* is the subject of the RC and the RC is the object of the main clause, which is unacceptable, and an overt accusative pronoun, here *den*, is needed as the antecedent, as shown in (8c).

(8)

- a. **Wer** das behauptet, ist ein Lügner.
who.NOM that claims, is a liar.
'(Anyone) who claims that is a liar.'
- b. Ich küsse, **wen** ich liebe.
I kiss, who.ACC I love.
'I kiss whoever I love.'
- c. **Wer** mir so etwas antut,
who.NOM to.me such a thing does,
den schaue ich nicht mehr an.
him.ACC look I no more at.
'I won't look at (anyone) who does something like that to me.'
- d. ***Wer** mir so etwas antut, schaue ich nicht mehr an.

The data in (8) confirm the analysis proposed by Tesnière (1959) that a *wh*-word occupies a double syntactic position (Kahane, 2002; Kahane and Osborne, 2015) and is both the head of the RC and a pronoun governed in the RC. We argue that the aim of the UD tree is not to reflect the exact structure, and that it is reasonable to attribute a unique

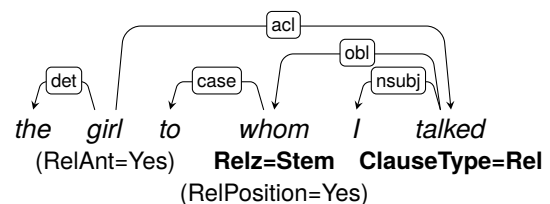
position to the relative pronoun to enhance comparison. The solution we propose in figure 7 has the advantage to maintain similar analyses for relative pronouns in RCs with or without antecedent. Following the same principles, we propose a different analysis for free RCs in languages with other types of relativizers, such as Wolof in section 6.

4. Head-External Relative Clauses

We now come back to our first examples, which were head-external RCs (HERCs), that is, RCs where there is a noun or a pronoun, called the antecedent, that governs the RC and is coreferent with the relativized position. This is by far the most common strategy across European languages and it has often been analyzed as an extraction of the relativized element. Our analysis of HIRCs can be easily extended to our two examples of HERCs in (1), giving the analyses in (9). The sub-relation *relcl* is replaced by the feature *ClauseType=Rel* and the feature *PronType=Rel* by *Relz=Stem*.¹⁰ A feature *RelPosition=Yes* has been added on the resumptive pronoun in (9b). We put into brackets the features that are optional because they are easily deductible from the structure.

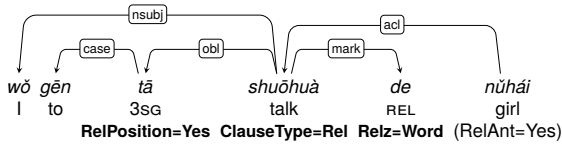
(9)

a. English



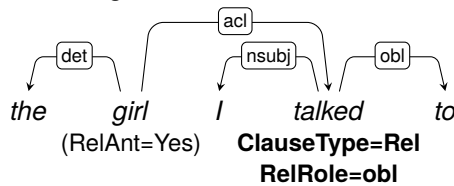
b. Mandarin

¹⁰It is also possible to keep *PronType=Rel* for relative pronouns, especially for treebanks that associate a feature *PronType* to every pronoun. Note that here the relative pronoun is inflected with a case suffix and the relativizer itself is the stem.

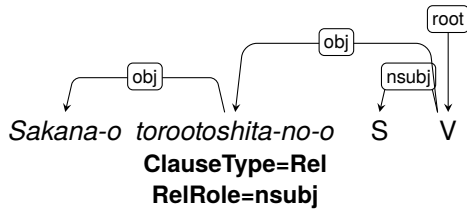


English also uses the gap strategy as in (10a), where the argument that is relativized in the RC is not expressed. Without a relative pronoun, it is not possible to recover information about the syntactic role of the relativized element from the annotation. To be able to quantify this information and allow more fine-grained cross-linguistic studies, we propose to introduce the feature `RelRole` on the governor of the gap with its syntactic role as a value. The same analysis can be applied to the free RC of Japanese we presented in (6).

(10) a. English



b. Japanese



The current UD annotation guidelines proposed to encode the position to the gap by an enhanced dependency from the governor of the gap to the antecedent (Universal Dependencies, 2026b). But this solution would not be easy for the free RC of Japanese in (10b), because there is no antecedent. Moreover, the enhanced dependency level uses a different encoding of the dependencies and only a few treebanks have implemented it. Here we just propose to introduce a `RelRole` feature that can be used to easily produce a correct enhanced level of annotation.

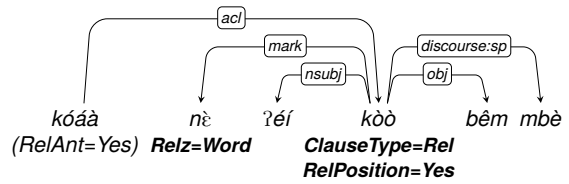
Another interesting example is that of RCs which are both externally and internally headed. We propose to call them *RCs with a resumptive antecedent*.¹¹ A puzzling example is found in some Gbaya RCs, such as (11), where there is a nominalized verbal antecedent that is also present in the RC in a finite form. Since the relative clause overtly expresses all arguments, it is the verb that is coreferent to the head of the RC. To mark this, we again use the feature `RelPosition=Yes`.

¹¹Hu et al. (2016) describe RCs "with a resumptive NP" in Mandarin children's productions.

(11) Gbaya (v2.17, Roulon-Doko et al., 2025)

kóáà nè ʔéí kòò
 give_birth.NMLZ REL LOG.SG give_birth.PFV.MT
 bêm mbè zók mò ná
 child amazing see.IPFV thing DISC
 ká bêmí ʔá-nè dèè
 and_then child.ANAPH TOP do.PFV.MT
 déŋ-kéyéŋ ʔòóyé
 tall_(6_8_years_old) of_course

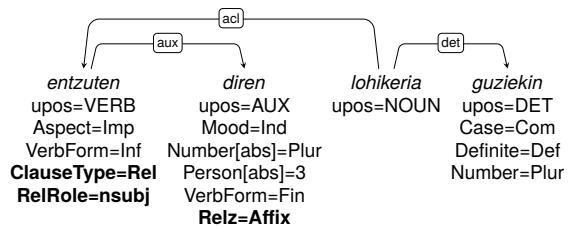
'Having surprisingly given birth to a child, she realizes that the child is of course grown up [6-8 years].'



5. Relative Clauses with Relativizer on the Verb

Usually, the relativizer is a verbal affix rather than a separate form. We have already presented the case of Japanese in (3c). This is the only strategy in various languages. In Basque, the relativization is ensured by a suffix `-(e)n` on the finite form (which is generally an auxiliary). For example, in (12), the finite auxiliary *diren* (lemma: *izan*) is relativized by the suffix `-n`. We mark the type of clause and the relativized position with `RelRole=nsubj` on the verb and the presence of the relativizer with `Relz=Affix` on the auxiliary.

(12) Basque (v2.17, Aranzabe et al., 2015)



'with all nonsense that are heard'

In Turkish, the relativization is ensured by the participial form of the verb and the subject must be realized as a possessive nominal complement. In (13b), the nominal complement (`nmod`) *çocuk* 'child' of (13a) is relativized. The verb is now at the participial form *çalış-tığ* and the 1PL subject pronominal index `-uz` has been replaced by the possessive index `-ımız`. The role of the gap in the RC is indicated by the feature `RelRole=nmod` on the governor of the relativized position.

(13) Turkish (Creissels, 2006)

- a. O çösuğ-un baba-sı ile
DEM child-GEN father-CSTR with
çalış-ıyor-uz.
work-PROG-S1PL
'We work with this child's father.'
- b. [baba-sı ile çalış-tığ-ımız]
[father-CSTR with work-PART-POSS1PL]
RelRole=nmod **ClauseType=Rel**
Relz=Affix

çocuk
child
(RelAnt=Yes)

'the child whose father we work with',
lit. [our working with (the) father] child

Traditional grammars of European languages distinguish the RCs from participial clauses. Although the two constructions are different, there is a continuum between them, as shown by Basque and Turkish RCs. In both constructions, the clause's governor is co-referent with a position within the clause. In (14a) and (14b), we analyze English participial clauses. We use the feature `ClauseType=Rel`, but it would be possible to introduce a feature `ClauseType=Part`, assuming that we are able to draw a line between RCs and participial clauses and to decide on which side are Basque and Turkish. These two constructions share, with the construction in (14c), the fact that a position is missing in the subordinated clause. We consider them to be particular cases of non-finite RCs involving a gap (which role is indicated by the feature `RelRole`). These constructions must be clearly distinguished from (14d), where there is no gap in the subordinated clause. We assigned to this clause the feature `ClauseType=Comp`.

(14) English

- a. *people living in glass houses*
Relz=Affix
RelRole=nsubj
ClauseType=Rel
- b. *students given high marks*
Relz=Affix
RelRole=nsubj:pass
ClauseType=Rel
- c. *I found a house to live in*
Relz=Word
RelRole=obl
ClauseType=Rel
- d. *the decision to leave was difficult*
ClauseType=Comp

6. Relative Clauses with a Separate Relativizer

We have seen that the use of a relativizer as part of the verbal form is quite common, even in Indo-European languages, and that UD does not foresee the use of a specific relative feature on verbs. This is part of a general preference for pronominal analysis of relatives, while the word class of relativizer is more of a gradient continuum than a discrete category.

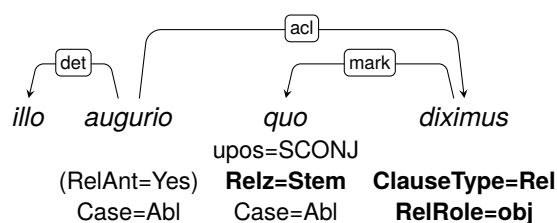
One strong argument for classifying a marker as a relative pronoun is its ability to agree with the antecedent while independently marking the case of the relativized element. In standard Latin (15a), the relative pronoun agrees in number and gender with the relative's head, but distinctly marks the accusative case of the extracted position (the object of the relative clause). However, Latin also exhibits case attraction, where the pronoun marks the case of the antecedent rather than retaining the case of the relativized position (15b) (Comrie, 1989; Creissels, 2006). This phenomenon blurs categorisation. While the relativizer retains its agreement properties, it behaves more like a modifying element that agrees with the head than as an independent argument of the clause. This is a common pattern across different languages.

(15) Latin (Creissels, 2006)

- a. illo augurio quod
DEM.SGN.ABL omen.SG.ABL RELZ.SGN.ACC
diximus
say.ACP.S3SG
'by that omen which we mentioned'
- b. illo augurio quo
DEM.SGN.ABL omen.SG.ABL RELZ.SGN.ABL
diximus
say.ACP.S3SG
'by that omen which we mentioned'

We think, with Creissels (2006), that this is where we must draw the line between relative pronouns and other relativizers. In (15b), the relativizer *quo* no longer occupies a position inside the relative clause, but it agrees with the antecedent. We call it an *agreeing relativizer*.

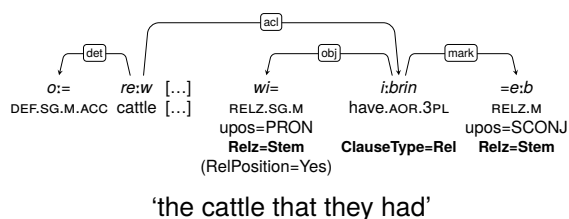
(16) Latin



Agreeing relativizers must be analyzed as markers of the relativization with relation `mark`, as in (16). For the POS, we use `SCONJ`, even if `PRON` would make sense too.

The contrasts between relative pronouns and agreeing relativizers is also well illustrated in Beja (Vanhove, 2014; Kahane et al., 2021). Beja uses a complex strategy where both an agreeing relativizer and a relative pronoun can coexist within the same structure, as in (17). While the agreeing relativizer marks the right boundary (Beja is a head-final language), the relative pronoun occupies the relativized position. Consequently, both elements may be annotated with the feature `Relz=Yes`, with the distinction between them relying only on their part of speech and syntactic roles.

(17) Beja (v2.17, Kahane et al., 2021)



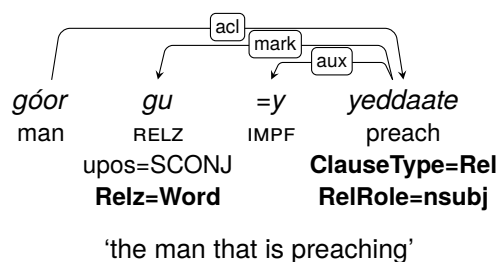
Following our analysis of relativizers in Latin and Beja, we think that, in many other languages, relativizers should be analyzed as agreeing relativizers (`upos=SCONJ`, `mark`) rather than relative pronouns. It is for instance the case of Wolof where the relativizer agrees in class noun with its antecedent, but does not mark the relativized position. In (18), the relativizer *gu* is the combination of a noun class marker *g-* (corresponding to the class of the antecedent noun) and a suffix *-u* (one of the three possible suffixes for relativizers).

The switch from the current annotation to this new annotation would not lose any information as soon as we have a feature such as `RelRole` to mark the syntactic role of the relativized position.

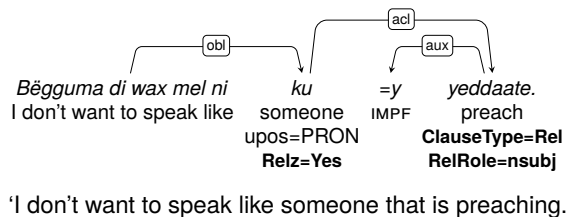
Wolof has also a lot of “free” RCs, where the noun phrase starts by a relativizer. These relativizers have the same form as for RCs with nominal antecedent, but in this case the noun class is meaningful. For instance, in (18b), the *k-* of the relativizer *ku* means ‘someone’ and the relativizer must be analyzed as a pronoun. Contrary to the case of relative pronoun that occupy a position inside the RC (in addition to their head position), this pronoun does not mark the relativized position and must be analyzed as the head, as in the current UD analysis. Note that, with our analysis, the structure of the RC is very similar in (18) and (18b); in both cases, we have a gap and a `RelRole` feature indicating its position. This contrasts with the current UD analysis where the relativizer would be analyzed as filling the relativized position in example (18) but not in (18b).

(18) Wolof (v2.17, Dione, 2019)

a. Free RC

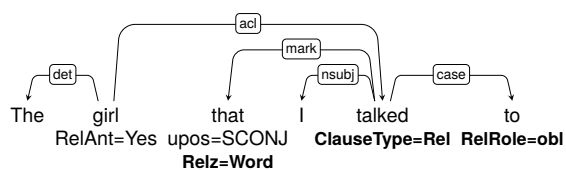


b. RC with pronominal antecedent



The third case of relativizers are pure relativizers, that do not occupy the relativized position in the RC and do not agree with the antecedent. This was illustrated by our Mandarin example (9b). We can also mention languages, like English, where the categorial status of relativizers is highly controversial (in typology, see e.g., Comrie, 1989; Andrews, 2007; Croft, 2022). In (19), the marker *that* is invariant: it shows no agreement with the antecedent and serves primarily to mark the left boundary of the relative clause. Importantly, unlike true relative pronouns (such as *which*), *that* cannot keep the preposition with it (e.g., *in which* vs. **in that*, the so-called pied-piping).

(19) English



One might hesitate to label *that* as a pure relativizer (`SCONJ`), concerned that information about the role of the extracted element would be lost. However, the features introduced in this article preserve this structural information specifically. Rather than obscuring the data, this approach enables robust cross-linguistic comparisons to be made: constructions with highly specific, language-dependent properties can now be analyzed using a unified annotation scheme. This makes the datasets comparable and explicitly encodes the typological differences, allowing them to be systematically quantified.

	Strategy	Relz	RC Features	Example
HIRC	Correlative	upos=DET PRON	RelPosition, RelAnt	Bambara
	Prototypical HIRC	Verbal affix	RelPosition	Japanese
	Free relative	upos=PRON upos=SCONJ Verbal affix	RelPosition RelRole RelRole	English, German Wolof Japanese
	Resumptive antecedent	upos=SCONJ	RelPosition	Gbaya
	Resumptive pronoun	upos=SCONJ	RelPosition	Mandarin
HERC	Gap	upos=SCONJ Verbal affix ∅	RelRole	English, Wolof Basque, Turkish, English participial English
	Relative pronoun	upos=PRON upos=PRON + upos=SCONJ	RelPosition RelPosition	English, Latin Beja

Table 1: Overview of the RC strategies analyzed and their encoding features. The feature `ClauseType=Rel` is not included as it is always annotated.

7. Conclusion

The goal of this paper was to highlight certain limitations in the current UD encoding of relative clauses and to propose a more robust encoding capable of handling a wide range of cross-linguistic strategies. Although we do not claim to cover every possible strategy, our review provides a broad foundation, and we believe the proposed encoding is sufficiently flexible to annotate additional strategies beyond those explicitly discussed. Ultimately, the encoding must be discussed with the entire UD community to determine its viability and explore whether an equivalent encoding can be developed within UD, or through another constructional approach if not (e.g., [Weissweiler et al., 2024](#)).

It should be noted that the conversion from the current UD encoding to the proposed encoding can quite easily be done with graph rewriting system such as Grew ([Guillaume, 2021](#)). But for many strategies of relativization, current UD treebanks lack information, especially when the relativizer is a verbal affix or the relativized position is a gap or a resumptive pronoun, which is likely to be the case for many languages in the UD collection. Moreover, there are currently no clear lines drawn between relative pronouns and other relativizers and many analyses should be revised to achieve a more homogeneous collection of treebanks, even for well-described languages, such as English or French ([Kayne, 1974](#); [Kahane, 2002](#)), where it has been shown that they mix relative pronouns and pure relativizers.

Our new annotation is based on features identifying the RC and its four types of element: relativizers, the relativized position, the role of the relativized position (when it is a gap or a resumptive pronoun or antecedent), and the antecedent in the main clause (when it is not the governor of the clause). In table 1, we propose a summary of the

various strategies of relativization presented here and the features used to encode them. Each time, we indicate the nature of the relativizer(s).

The decision to use clause-level features is non-trivial and creates natural tension within a dependency framework. However, the debate over incorporating constituency-oriented annotations is already occurring in practice and is proving essential for cross-linguistic comparison. Throughout this paper, we have outlined some theoretical and practical advantages of using these features rather than subtype relations. Interpreting these features as scoped to a node’s projection is straightforward, as long as the scope of each feature is clear. This requires a clear categorisation of features by their structural level (e.g., lexical vs clausal). Ultimately, the use of these features should also be a topic of discussion within UD.

Finally, another advantage of our encoding is that it can be extended to other constructions, such as interrogation and clefting (20). While interrogation is structurally (and genetically) similar to relativization, it is currently not annotated in UD treebanks. And in clefting, the role of the extracted element is currently not indicated (unless there is an enhanced level of analysis).

- (20) a. Relativization
the girl who biked
upos=PRON ClauseType=Rel
Relz=Word
- b. Interrogation
I wonder who biked
upos=PRON ClauseType=Int
PronType=Int
- c. Clefting
It is her that biked.
upos=SCONJ ClauseType=Cleft
CleftRole=nsubj

8. Acknowledgements

We would like to thank our colleagues for the discussions we have had about the data: Katya Aplonova for Bambara, Martine Vanhove for Beja, and Qishen Wu for Mandarin.

9. Bibliographical References

- Avery D. Andrews. 2007. Relative clauses. *Language Typology and Syntactic Description*, 2:206–236.
- Ekaterina Aplonova and Francis M. Tyers. 2017. [Towards a dependency-annotated treebank for Bambara](#). In *Proceedings of the 16th International Workshop on Treebanks and Linguistic Theories*, pages 138–145, Prague, Czech Republic. Association for Computational Linguistics.
- Maxux Aranzabe, Aitziber Atutxa, Kepa Bengoetxea, Arantza Díaz de Ilarraza, Iakes Goenaga, Koldo Gojenola, and Larraitz Uribe. 2015. Automatic conversion of the Basque dependency treebank to Universal Dependencies. In *Proceedings of the 14th International Workshop on Treebanks and Linguistic Theories (TLT 2015)*, pages 233–241, Warsaw, Poland. Institute of Computer Science of the Polish Academy of Sciences.
- David Basilico. 1996. [Head position and internally headed relative clauses](#). *Language*, 72(3):498–532.
- Joan Bresnan and Jane Grimshaw. 1978. The syntax of free relatives in English. *Linguistic Inquiry*, 9(3):331–391.
- Bernard Comrie. 1989. *Relative Clauses*, 2nd edition, chapter 7. University of Chicago Press, Chicago.
- Bernard Comrie. 1998. Rethinking the typology of relative clauses. *Language Design: journal of theoretical and experimental linguistics*, 1:059–85.
- Denis Creissels. 2006. *Relatives postnominales*, chapter 33. Lavoisier, Paris.
- William Croft. 2022. [Events as Modifiers: Relative Clause Constructions](#), chapter 19. Cambridge University Press, Cambridge.
- William Croft and Joakim Nivre. 2025. [Verbal predication constructions in Universal Dependencies](#). In *Proceedings of the Second International Workshop on Construction Grammars and NLP*, pages 50–60, Düsseldorf, Germany. Association for Computational Linguistics.
- William Croft, Dawn Nordquist, Katherine Looney, and Michael Regan. 2017. Linguistic typology meets Universal Dependencies. In *Proceedings of the 15th International Workshop on Treebanks and Linguistic Theories (TLT15)*, pages 63–75. CEUR Workshop Proceedings.
- Christopher D. Culy. 1990. *The syntax and semantics of internally headed relative clauses*. Stanford University.
- Cheikh Bamba Dione. 2019. [Developing Universal Dependencies for Wolof](#). In *Proceedings of the Third Workshop on Universal Dependencies (UDW, SyntaxFest)*, pages 12–23, Paris, France. Association for Computational Linguistics.
- Bruce T. Downing. 1978. Some universals of relative clause structure. In Joseph H. Greenberg, editor, *Universals of Human Language, Volume 4: Syntax*, pages 375–418. Stanford University Press, Stanford, CA.
- Kim Gerdes, Sylvain Kahane, and Xinying Chen. 2021. [Typometrics: From Implicational to Quantitative Universals in Word Order Typology](#). *Glossa: a journal of general linguistics*, 6(1):17.
- Larry P/ Gorbet. 1974. *Relativization and complementation in Diegueno: Noun phrases as nouns*. Ph.D. thesis, University of California, San Diego.
- Alexander Grosu and Koji Hoshi. 2016. Japanese internally headed relatives: Their distinctness from potentially homophonous constructions. *Glossa: a journal of general linguistics*, 1(1).
- Alexander Grosu and Fred Landman. 1998. Strange relatives of the third kind. *Natural Language Semantics*, 6(2):125–170.
- Bruno Guillaume. 2021. [Graph matching and graph rewriting: GREW tools for corpus exploration, maintenance and conversion](#). In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations*, pages 168–175, Online. Association for Computational Linguistics.
- Shenai Hu, Anna Gavarró, and Maria Teresa Guasti. 2016. Children’s production of head-final relative clauses: The case of Mandarin. *Applied Psycholinguistics*, 37(2):323–346.
- Sylvain Kahane. 2002. À propos de la position syntaxique des mots qu-. *Verbum (Presses Universitaires de Nancy)*, 24(4).

- Sylvain Kahane and Timothy F. Osborne. 2015. Translators' introduction. In *Lucien Tesnière, Elements of structural syntax*, pages ixxx–lxiii. John Benjamins.
- Sylvain Kahane, Martine Vanhove, Rayan Ziane, and Bruno Guillaume. 2021. [A morph-based and a word-based treebank for Beja](#). In *Proceedings of the 20th International Workshop on Treebanks and Linguistic Theories (TLT, SyntaxFest 2021)*, pages 48–60, Sofia, Bulgaria. Association for Computational Linguistics.
- Amanda Kann. 2024. [Massively multilingual token-based typology using the parallel Bible corpus](#). In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING)*, pages 11070–11079, Torino, Italia. ELRA and ICCL.
- Richard Kayne. 1974. French relative *que* (1). *Recherches Linguistiques*, 11:40–61.
- Edward L. Keenan and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8(1):63–99.
- Sige-Yuki Kuroda. 1974. Pivot-independent relativization in Japanese (i). *Journal of Japanese Linguistics*, 3(1-2):61–96.
- Pierre Le Goffic. 2002. Marqueurs d'interrogation/indéfinition/subordination: essai de vue d'ensemble. *Verbum*, 24(4):315–340.
- Christian Lehmann. 1986. [On the typology of relative clauses](#). *Linguistics*, 24(4):663–680.
- Natalia Levshina. 2019. [Token-based typology and word order entropy: A study based on universal dependencies](#). *Linguistic Typology*, 23(3):533–572.
- Natalia Levshina. 2022. [Corpus-based typology: applications, challenges and some solutions](#). *Linguistic Typology*, 26(1):129–160.
- Keiko Murasugi. 1991. *Noun phrases in Japanese and English: A study in syntax, learnability and acquisition*. Ph.D. thesis, University of Connecticut.
- Joakim Nivre. 2025. [Constructions and strategies in Universal Dependencies](#). In *Proceedings of the Joint 25th Nordic Conference on Computational Linguistics and 11th Baltic Conference on Human Language Technologies (NoDaLiDa/Baltic-HLT 2025)*, pages 419–423, Tallinn, Estonia. University of Tartu Library.
- Joakim Nivre and William Croft. 2025. [Reference and modification in Universal Dependencies](#). In *Proceedings of the Eighth Workshop on Universal Dependencies (UDW, SyntaxFest 2025)*, pages 1–10, Ljubljana, Slovenia. Association for Computational Linguistics.
- Robert Östling. 2015. [Word order typology through multilingual word alignment](#). In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 205–211, Beijing, China. Association for Computational Linguistics.
- Paul R. Platero. 1973. A study of the Navajo relative clause. Master's thesis, Massachusetts Institute of Technology.
- Paulette Roulon-Doko, Sylvain Kahane, and Bruno Guillaume. 2025. [A morpheme-based treebank for Gbaya, an Ubangian language of Central Africa](#). In *Proceedings of the Eighth International Conference on Dependency Linguistics (Depling, SyntaxFest 2025)*, pages 93–102, Ljubljana, Slovenia. Association for Computational Linguistics.
- Natalia Silveira, Timothy Dozat, Marie-Catherine de Marneffe, Samuel Bowman, Miriam Connor, John Bauer, and Chris Manning. 2014. [A gold standard dependency corpus for English](#). In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, pages 2897–2904, Reykjavik, Iceland. European Language Resources Association (ELRA).
- Lucien Tesnière. 1959. *Éléments de syntaxe structurale*. Klincksieck.
- Universal Dependencies. 2026a. [acl:crel: Correlative Clausal Modifier](#). <https://universaldependencies.org/sa/dep/acl-crel>. Accessed: January 2026.
- Universal Dependencies. 2026b. [Enhanced Dependencies: Relatives Clauses](#). <https://universaldependencies.org/u/overview/enhanced-syntax.html#relative-clauses>. Accessed: January 2026.
- Universal Dependencies. 2026c. [Pron-type: Pronominal Type](#). <https://universaldependencies.org/u/feat/PronType>. Accessed: January 2026.
- Universal Dependencies. 2026d. [Relative clause](#). <https://universaldependencies.org/>

[workgroups/newdoc/relative_clauses](#).
Accessed: January 2026.

Martine Vanhove. 2014. Beja grammatical sketch. *The corpAfroAs corpus of spoken AfroAsiatic languages*, pages 1–68.

Ralf Vogel. 2003. Surface matters. case conflict in the free relative constructions and case theory. In E. Brandner and H. Zinsmeister, editors, *New perspectives in Case Theory*. CSLI.

Leonie Weissweiler, Nina Böbel, Kirian Guiller, Santiago Herrera, Wesley Scivetti, Arthur Lorenzi, Nurit Melnik, Archana Bhatia, Hinrich Schütze, Lori Levin, Amir Zeldes, Joakim Nivre, William Croft, and Nathan Schneider. 2024. [UCxn: Typologically informed annotation of constructions atop Universal Dependencies](#). In *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING)*, pages 16919–16932, Torino, Italia. ELRA and ICCL.

10. Language Resource References

Daniel Zeman, Joakim Nivre, Mitchell Abrams, Elia Ackermann, Jephthé Adolphe, Noëmi Aeppli, Hamid Aghaei, Željko Agić, Amir Ahmadi, Lars Ahrenberg, Chika Kennedy Ajede, Arofat Akhundjanova, Furkan Akkurt, Gabrielè Aleksandravičiūtė, Ika Alfina, Avner Algom, Khalid Alnajjar, Chiara Alzetta, Antonios Anastasopoulos, Erik Andersen, Kirk Andrews, Matthew Andrews, Lene Antonsen, Tatsuya Aoyama, Katya Aplonova, Angelina Aquino, Carolina Aragon, Glyd Aranes, Maria Jesus Aranzabe, Bilge Nas Arcan, Hórunn Arnardóttir, Wirote Aroonmanakun, Gashaw Arutie, Jessica Naraiswari Arwidarasti, Hiwa Asadpour, Masayuki Asahara, Katla Ásgeirsdóttir, Deniz Baran Aslan, Cengiz Asmaoğlu, Luma Ateyah, Furkan Atmaca, Mohammed Attia, Aitziber Atutxa, Liesbeth Augustinus, Mariana Avelãs, Elena Badmaeva, Jana Bajorat, Keerthana Balasubramani, Miguel Ballesteros, Esha Banerjee, Sebastian Bank, Bryan Khelven da Silva Barbosa, Verginica Barbu Mititelu, Starkaður Barkarson, Rodolfo Basile, Victoria Basmov, Colin Batchelor, John Bauer, Seyyit Talha Bedir, Shabnam Behzad, Nathanaël Beiner, Juan Belieni, Alevtina Bémová, Kepa Bengoetxea, İbrahim Benli, Yifat Ben Moshe, Marie Benzerrak, Aleksandrs Berdicevskis, Ansu Berg, Gözde Berk, Delphine

Bernhard, Astrid Berntsson Ingelstam, Riyaz Ahmad Bhat, Erica Biagetti, Eckhard Bick, Agnė Bielinskienė, Esma Fatıma Bilgin Taşdemir, Helin Binici, Kristín Bjarnadóttir, Verena Blaschke, Rogier Blokland, Nina Böbel, Victoria Bobicev, Loïc Boizou, Stavros Bompolas, Johnatan Bonilla, Emanuel Borges Völker, Lars Borin, Carl Börstell, Cristina Bosco, Gosse Bouma, Sam Bowman, Adriane Boyd, Anouck Braggaa, António Branco, Myriam Bras, Théo Brillet, Kristina Brokaitė, Lanni Bu, Eva Buráňová, Aljoscha Burchardt, Carmen Cabeza, Natalia Cáceres Arandia, Olesea Caftanator, Marisa Campos, Marie Candito, Caterina Maria Cappello, Bernard Caron, Gauthier Caron, Catarina Carvalheiro, Rita Carvalho, Lauren Cassidy, Maria Clara Castro, Sérgio Castro, Tatiana Cavalcanti, Gülşen Cebiroğlu Eryiğit, Flavio Massimiliano Cecchini, Giuseppe G. A. Celano, Anila Çepani, Slavomír Čéplö, Neslihan Cesur, Savas Cetin, Özlem Çetinoğlu, Fabricio Chalub, Liyanage Chamila, Claudine Chamoreau, Shweta Chauhan, Yifei Chen, Ethan Chi, Taishi Chika, Yongseok Cho, Jinho Choi, Bermet Chontaeva, Jayeol Chun, Juyeon Chung, Alessandra T. Cignarella, Silvie Cinková, Esther Cocco, Aurélie Collomb, Çağrı Çöltekin, Miriam Connor, Claudia Corbetta, Daniela Corbetta, Francisco Costa, Marine Courtin, Benoît Crabbé, Mihaela Cristescu, Vladimir Cvetkoski, Netanel Dahan, Ingerid Løyning Dale, Sabrina D'Alì, Philemon Daniel, Khensa Daoudi, Bijayalaxmi Dash, Satya Ranjan Dash, Elizabeth Davidson, Leonel Figueiredo de Alencar, Mathieu Dehouck, Martina de Laurentiis, Marie-Catherine de Marneffe, Ahmet Demir, Valeria de Paiva, Mehmet Oguz Derin, Elvis de Souza, Arantza Diaz de Ilarraza, Roberto Antonio Díaz Hernández, Carly Dickerson, Ariani Di Felippo, Arawinda Dinakaramani, Elisa Di Nuovo, Bamba Dione, Peter Dirix, Hoa Do, Kaja Dobrovoljc, Caroline Döhmer, Adrian Doyle, Timothy Dozat, Kira Droganova, Magali Sanches Duran, Puneet Dwivedi, Christian Ebert, Hanne Eckhoff, Masaki Eguchi, Sandra Eiche, Roald Eiselen, Marhaba Eli, Ali Elkahky, Binyam Ephrem, Olga Erina, Tomaz Erjavec, Louise Esher, Soudabeh Eslami, Farah Essaidi, Aline Etienne, Wograine Evelyn, Sidney Facundes, Richárd Farkas, Ján Faryad, Federica Favero, Jannatul Ferdaousi, Marília Fernanda, Hector Fernandez Alcalde, Amal Fethi, Jennifer Foster, Barbara Francioni, Theodorus Fransen, Cláudia Freitas, Kazunori Fujita, Katarína Gajdošová, Daniel Galbraith, Edith Galy, Federica Gamba, Marcos Garcia, José María García-Miguel, Moa Gärdenfors, Tanja Gaus-tad, Efe Eren Genç, Fabrício Ferraz Ger-

ardi, Kim Gerdes, Luke Gessler, Filip Ginter, Gustavo Godoy, Iakes Goenaga, Koldo Gojenola, Memduh Gökirmak, Yoav Goldberg, Gili Goldin, Xavier Gómez Guinovart, Berta González Saavedra, Mathieu Goux, Bernadeta Griciūtė, Matias Grioni, Loïc Grobol, Normunds Grūzītis, Mario Guglielmetti, Bruno Guillaume, Kirian Guillier, Céline Guillot-Barbance, Tunga Güngör, Vladimir Gurevich, Nizar Habash, Henrik Hafsteinsson, Michael Hahn, Jan Hajič, Jan Hajič jr., Eva Hajičová, Mika Hämäläinen, Linh Hà Mỹ, Na-Rae Han, Muhammad Yudistira Hanifmuti, Takahiro Harada, Sam Hardwick, Kim Harris, Naïma Hassert, Dag Haug, Jiří Havelka, Johannes Heinecke, Oliver Hellwig, Felix Hennig, Barbora Hladká, Jaroslava Hlaváčová, Florinel Hociung, Diana Hoefels, Barbara Hoff, Petter Hohle, Nick Howell, Yidi Huang, Marivel Huerta Mendez, Jena Hwang, Takumi Ikeda, Inessa Iliadou, Anton Karl Ingason, Radu Ion, Elena Irimia, Olájiḍé Ishola, Artan Islamaj, Kaoru Ito, Federica Iurescia, Jessica K. Ivani, Sandra Jagodzińska, Siratun Jannat, Tomáš Jelínek, Apoorva Jha, Ratanon Jiamsundutsadee, Katharine Jiang, Sylvanus Job, Mayank Jobanputra, Anders Johannsen, Hildur Jónsdóttir, Fredrik Jørgensen, Zhuoxuan Ju, Markus Juutinen, Hüner Kaşıkara, Nadezhda Kabaeva, Sylvain Kahane, Hiroshi Kanayama, Jenna Kanerva, Neslihan Kara, Ritván Karahóga, Jiří Kárník, Andre Kåsen, Tolga Kayadelen, Sarveswaran Kengatharayer, Václava Kettnerová, Lilit Kharatyan, Jesse Kirchner, Elena Klementieva, Elena Klyachko, Petr Kocharov, Arne Köhn, Abdullatif Köksal, Veronika Kolářová, Kamil Kopacewicz, Timo Korkiakangas, Mehmet Köse, Alexey Koshevoy, Nelda Kote, Natalia Kotsyba, Barbara Kovačić, Jolanta Kovalevskaitė, Emmanuelle Kowner, Simon Krek, Parameswari Krishnamurthy, Sandra Kübler, Lucie Kučová, Adrian Kuqi, Elmurod Kuriyozov, Oğuzhan Kuyrukçu, Asli Kuzgun, Sookyung Kwak, Kris Kyle, Käbi Laan, Veronika Laippala, Lorenzo Lambertino, Israel Landau, Tatiana Lando, Septina Dian Larasati, Pierre Larrivéé, Kusum Lata, Alexei Lavrentiev, John Lee, Phương Lê Hồng, Alessandro Lenci, Wei Qi Leong, Saran Lertpradit, Herman Leung, Lori Levin, Maria Levina, Lauren Levine, Cheuk Ying Li, Josie Li, Keying Li, Yixuan Li, Yuan Li, KyungTae Lim, Bruna Lima Padovani, Yi-Ju Jessica Lin, Krister Lindén, Yang Janet Liu, Zoey Liu, Nikola Ljubešić, Irina Lobzhanidze, Olga Loginova, Markéta Lopatková, Lucelene Lopes, Edita Luftiu, Arsenii Lukashevskiy, Stefano Lusito, Anne-Marie Lutgen, Andry Luthfi, Mikko Luukko, Olga Lyashevskaya, Teresa Lynn, Vivien Macketanz, Menel Mahamdi,

Jean Maillard, Punyanuch Maitreenukul, Ilya Makarchuk, Aibek Makazhanov, Francesco Mambrini, Michael Mandl, Christopher Manning, Ruli Manurung, Büşra Marşan, Cătălina Mărânduc, David Mareček, Katrin Marheinecke, Stella Markantonatou, Héctor Martínez Alonso, Lorena Martín Rodríguez, André Martins, Cláudia Martins, Arianna Masciolini, Jan Mašek, Sanatbek Matlatipov, Hiroshi Matsuda, Yuji Matsumoto, Caterina Mauri, Alessandro Mazzei, Ryan McDonald, Sarah McGuinness, Maitrey Mehta, Pierre André Ménard, Gustavo Mendonça, Hilla Merhav, Tatiana Merzhevich, Paul Meurer, Niko Miekka, Marie Mikulová, Emilia Milano, Aleksandra Miletić, Aaron Miller, Junghyun Min, Yael Minerbi, Jiří Mirovský, Karina Mischenkova, Anna Missilä, Cătălin Mititelu, Maria Mitrofan, Yusuke Miyao, Biswakalpita Mohapatra, AmirHossein Mojiri Foroushani, Judit Molnár, Amirsaeid Moloodi, Simonetta Montemagni, Amir More, Laura Moreno Romero, Giovanni Moretti, Shinsuke Mori, Tomohiko Morioka, Shigeki Moro, Bjartur Mortensen, Bohdan Moskalevskiy, Katerina Mouzou, Kadri Muischnek, Robert Munro, Yugo Murawaki, Nikolett Mus, Kaili Müürisep, Pinkey Nainwani, Mariam Nakhlé, Mino Nassajian, Juan Ignacio Navarro Horñiácek, Anna Nedoluzhko, Gunta Nešpore-Bērzkalne, Manuela Nevaci, Lương Nguyễn Thị, Huyền Nguyễn Thị Minh, Yoshihiro Nikaido, Vitaly Nikolaev, Rattima Nitisaroj, Victor Norrman, Alireza Nourian, Michal Novák, Maria das Graças Volpe Nunes, Hanna Nurmi, Stina Ojala, Atul Kr. Ojha, Hulda Óladóttir, Adédayò Olúòkun, Mai Omura, Emeka Onwuegbuzia, Noam Ordan, Petya Osenova, Robert Östling, Annika Ott, Lilja Øvreliid, Masanori Oya, Şaziye Betül Özateş, Merve Özçelik, Arzucan Özgür, Balkız Öztürk Başaran, Teresa Paccosi, Petr Pajas, Thomas Palakapilly, Alessio Palmero Aprosio, Jarmila Panevová, Ludovica Pannitto, Anastasia Panova, Thiago Alexandre Salgueiro Pardo, Shantipriya Parida, Hyunji Hayley Park, Niko Partanen, Elena Pascual, Marco Passarotti, Agnieszka Patejuk, Guilherme Paulino-Passos, Giulia Pedonese, Oggi Peeters, Angelika Peljak-Łapińska, Siyao Peng, Siyao Logan Peng, Rita Pereira, Sílvia Pereira, Cenel-Augusto Perez, Natalia Perkova, Guy Perrier, Slav Petrov, Daria Petrova, Eva Pettersson, Andrea Peverelli, Jason Phelan, Claudel Pierre-Louis, Jussi Piitulainen, Yuval Pinter, Clara Pinto, Rodrigo Pintucci, Tommi A Pirinen, Emily Pitler, Magdalena Plamada, Barbara Plank, Alistair Plum, Thierry Poibeau, Charin Polpanumas, Larisa Ponomareva, Martin Popel, Clamença Poujade, Rangga Prangwedana Prangwedana, Lauma Pretkalniņa, Rigardt Preto-

rius, Sophie Prévost, Prokopis Prokopidis, Adam Przepiórkowski, Robert Pugh, Tiina Puolakainen, Christoph Purschke, Sampo Pyysalo, Peng Qi, Andreia Querido, Andriela Rääbis, Ella Rabinovich, Alexandre Rademaker, Mutee-u Rahman, Mizanur Rahoman, Taraka Rama, Loganathan Ramasamy, Carlos Ramisch, Joana Ramos, Fam Rashel, Mohammad Sadegh Rasooli, Vinit Ravishankar, Livy Real, Petru Rebeja, Siva Reddy, Mathilde Regnault, Georg Rehm, Arij Riabi, Ivan Riabov, Michael Riebler, Erika Rimkutė, Larissa Rinaldi, Laura Rituma, Putri Rizqiyah, Luisa Rocha, Eiríkur Rögnvaldsson, Ivan Roksandic, Norton Trevisan Roman, Mykhailo Romanenko, Natalia Romanova, Rudolf Rosa, Valentin Roşca, Paulette Roulon, Davide Rovati, Ben Rozonoyer, Olga Rudina, Jack Rueter, Paolo Ruffolo, Kristján Rúnarsson, Rozana Rushiti, Attapol T. Rutherford, Shoval Sadde, Pegah Safari, Aleksí Sahala, Kalyanamalini Sahoo, Saraswati Sahoo, Shadi Saleh, Alessio Salomoni, Tanja Samardžić, Konstantinos Sampanis, Stephanie Samson, Xulia Sánchez-Rodríguez, Manuela Sanguinetti, Ezgi Saniyar, Dage Särg, Marta Sartor, Albina Sarymsakova, Mitsuya Sasaki, Baiba Saulīte, Agata Savary, Yanin Sawanakunanon, Shefali Saxena, Kevin Scannell, Salvatore Scarlata, Emmanuel Schang, Robert Schikowski, Nathan Schneider, Sebastian Schuster, Lane Schwartz, Djamé Seddah, Wolfgang Seeker, Sven Sellmer, Mojgan Seraji, Magda Ševčíková, Petr Sgall, Syeda Shahzadi, Mo Shen, Atsuko Shimada, Gyu-Ho Shin, Hiroyuki Shirasu, Yana Shishkina, Muh Shohibussirri, Maria Shvedova, Jean Sibille, Janine Siewert, Einar Freyr Sigurðsson, João Silva, Aline Silveira, Natalia Silveira, Sara Silveira, Maria Simi, Radu Simionescu, Katalin Simkó, Mária Šimková, Haukur Barri Símonarson, Kiril Simov, Dmitri Sitchinava, Ted Sither, Aaron Smith, Isabela Soares-Bastos, Per Erik Solberg, Dolores Solberger, Barbara Sonnenhauser, Shafi Sourov, Nina Speransky, Rachele Sprugnoli, Panyut Sriwirote, Vivian Stamou, Steinhórf Steingrímsson, Antonio Stella, Jan Štěpánek, Barbora Štěpánková, Abishek Stephen, Milan Straka, Omer Strass, Emmett Strickland, Jana Strnadová, Sara Stymne, Alane Suhr, Yogi Lesmana Sulestio, Umut Sulubacak, Jack Sun, Hakyung Sung, Shingo Suzuki, Daniel Swanson, Zsolt Szántó, Maria Irena Szawerna, Chihiro Taguchi, Dima Taji, Luigi Talamo, Fabio Tamburini, Mary Ann C. Tan, Takaaki Tanaka, Dipta Tanaya, Mirko Tavoni, Nursena Teker, Samson Tella, Isabelle Tellier, Marinella Testori, Santhawat Thanyawong, Guillaume Thomas, Tarik Emre Tıraş, William Chandra Tjhi,

Thea Tollersrud, Kamil Tomaszek, Sara Tonelli, Liisi Torga, Lucas Toribio, Marsida Toska, Trond Trosterud, Anna Trukhina, Reut Tsarfaty, Kira Tulchynska, Utku Türk, Francis Tyers, Sveinbjörn Hórfarson, Vilhjálmur Hórfsteinsson, Sumire Uematsu, Roman Untilov, Zdeňka Urešová, Larraitz Uria, Hans Uszkoreit, Andrius Utkā, Elena Vagnoni, Sowmya Vajjala, Socrates Vak, Socrates Vakirtzian, Rob van der Goot, Martine Vanhove, Daniel van Niekerk, Gertjan van Noord, Viktor Varga, Helena Vaz, Uliana Vedenina, Giulia Venturi, Marianne Vergez-Couret, Annemarie Verkerk, Barbora Vidová Hladká, Eric Villemonte de la Clergerie, Veronika Vincze, Anishka Vissamsetty, Natalia Vlasova, Eleni Vligouridou, Aya Wakasa, Joel C. Wallenberg, Lars Wallin, Abigail Walsh, John Wang, Jonathan North Washington, Leonie Weissweiler, Maximilian Wendt, Paul Widmer, Aleksandra Wiecezorek, Shira Wigderson, Sri Hartati Wijono, Vanessa Berwanger Wille, Seyi Williams, Miriam Winkler, Shuly Wintner, Mats Wirén, Christian Wittern, Alena Witzlack-Makarevich, Tsegay Woldemariam, Tak-sum Wong, Alina Wróblewska, Qishen Wu, Mary Yako, Kayo Yamashita, Naoki Yamazaki, Chunxiao Yan, Qizhen Yang, Xiulin Yang, Koichi Yasuoka, Marat M. Yavrumyan, Arife Betül Yenice, Enes Yilandilođlu, Olcay Taner Yıldız, Zhuoran Yu, Arlisa Yuliawati, Zdeněk Žabokrtský, Shorouq Zahra, Amir Zeldes, Annie Zhang, Larry Zhang, He Zhou, Hanzhi Zhu, Yilun Zhu, Anna Zhuravleva, Rayan Ziane, Artūrs Znotiņš, and Eleonora Zucchini. 2025. [Universal dependencies 2.17](#). LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL).