Nonprototypical Predication and Nonpredicational Clauses in Universal Dependencies

Joakim Nivre, William Croft, André V. Lopes Coneglian
Uppsala University, University of New Mexico, Federal University of Minas Gerais
joakim.nivre@lingfil.uu.se, wacroft@icloud.com, coneglian@ufmg.br

Abstract

To assess whether the framework of Universal Dependencies (UD) is compatible with findings from linguistic typology, we need to systematically review how UD represents linguistic constructions and how it handles the range of morphosyntactic variation attested across languages. In this paper, we present such a review focusing on nonprototypical predication and nonpredicational clauses. We find that, while nonprototypical predication is generally handled well in the UD framework, nonpredicational clauses are not discussed as such in the guidelines and often have to be annotated in a way that does not reflect their special information packaging functions. We briefly discuss ways in which the UD framework could be extended in order to better capture these functions.

Keywords: Universal Dependencies, nonprototypical predication, nonpredicational clauses

1. Introduction

Universal Dependencies (UD) is an annotation framework for morphosyntax, designed to be applicable to all human languages and to enable meaningful cross-linguistic comparisons (Nivre et al., 2016, 2020; de Marneffe et al., 2021). To check if UD meets these requirements, Nivre (2025) proposes to build a constructicon for UD based on the survey of universal constructions and morphosyntactic strategies in Croft (2022) and the MoCCA database of comparative concepts derived from it (Lorenzi et al., 2024).

In this framework, *constructions* are form-function pairings defined solely in terms of their function (hence universal), while *strategies* are defined by the pairing of a function with some cross-linguistically identifiable morphosyntactic form. For example, the relative clause construction is defined as any construction in any language where an action concept modifies a referent. English *the book I'm reading* is an instance of the relative clause construction. The externally-headed strategy describes a relative clause construction in which the head noun is external to the modifying clause. The English example is an instance of this narrower category as well.

To build a constructicon for UD, we need to provide a UD analysis of every construction-strategy pair, and this paper lays the ground for such a resource by reviewing the UD guidelines for a specific set of constructions – nonprototypical predication and nonpredicational clauses – constructions that are discussed in Chapters 10–11 of Croft (2022). This complements earlier studies on reference and modification (Nivre and Croft, 2025) and verbal predication (Croft and Nivre, 2025), as well as concurrent studies on complex predicates (Coneglian et al., 2026) and speech act constructions (Nivre et al., 2026).

The first goal of the investigation is to check whether the UD guidelines cover the main attested strategies for each construction, and to provide new recommendations in case of gaps. The second goal is to find out whether the annotations of different strategies reveal that they are realizations of the same construction, which facilitates cross-linguistic comparison, and to discuss alternative analyses when this is not the case. We are aware that UD was not originally designed to capture constructions in Croft's sense, but trying to capture both constructions and strategies seems compatible with the view of UD as a "mixed functional-structural system" for annotation.[1] Finally, it is worth noting that our review focuses on the UD guidelines, not on the actual annotations in existing UD treebanks, although we occasionally make observations on treebanks when

[1] https://universaldependencies.org/u/overview/syntax.html#the-taxonomy-of-typed-dependencies

relevant. A systematic review of existing treebanks would be very valuable but is outside the scope of this project.

2. Nonprototypical Predication

Nonprototypical predication is the predication of concepts other than the action concepts associated with verbal clause constructions, previously discussed in Croft and Nivre (2025). This includes at least the following three types of predication (Croft, 2022, p. 289):

- Object predication: *Ann is an engineer*
- Property predication: *Sandy is clever*
- Locative predication: *the cat is on the mat*

There are four main morphosyntactic strategies used for nonprototypical predication in the world's languages. First, the *verbal* strategy consists in recruiting the morphosyntactic encoding of prototypical verbal predication for nonprototypical predication, as illustrated in example (1), from Moru (Tucker, 1940; Stassen, 1997), which compares (a) action predication and (b) property predication.

(1) a. anyá k-ɔ́dra
3PL 3PL.IMPF-die
'they are/were dying'

b. lédr aná k^{w}ɔ́izi
man that 3SG.IMPF-bad
'that man is bad'

The indexation on the predicate is identical in the two cases, even though the predicate stem in (b) may be used for nominal reference in other contexts.

Second, the *verbal copula* strategy combines the nonprototypical predicate with an independent word inflected like a verb. It is commonly used for locative predication, with a copula derived from a locational or posture verb, which can also be extended to other types of nonprototypical predication. This is exemplified in (2), from Amele (Roberts, 1987; Stassen, 1997), for (a) locative predication and (b) property predication.

(2) a. uqa jo na bil-i-a
he house at sit-3SG-PRS
'he is in the house'

b. uqa me bil-i-a
he good sit-3SG-PRS
'he is well'

Third, the *nonverbal copula strategy* instead combines the nonprototypical predicate with an uninflected form, most typically a pronoun (demonstrative or personal), a topic or focus marker, or some other grammaticalized form. Example (3) shows object predication with a nonverbal copula derived from a demonstrative in Nakanai (Stassen, 1997, p. 149), and example (4) shows object predication with a nonverbal copula derived from a focus marker in Awtuw (Johnston, 1980; Stassen, 1997).

(3) eia la taua sesele
3SG DEM spirit truly
'he is truly a spirit

(4) wan po rumeyæn
1SG FOC human-being
'I am a human being'

Fourth, the *zero strategy* simply juxtaposes the subject nominal with the nonprototypical predicate without any inflection or linking element. This is illustrated in (5), from Tiwi (Osborne, 1974; Stassen, 1997), for (a) object predication and (b) property predication.

(5) a. purukupaɪli maɪntina
parukuparli boss
'Parukuparli is boss'

b. tuŋkwaltiriŋa pumpuka
stringy-bark good
'a string bark is good'

Both the zero and the nonverbal copula strategies typically originate in object predication.

2.1. UD Annotation

The UD annotation of a nonprototypical predication construction connects the predicate to its argument with a nominal subject (*nsubj*) relation. This means that, as far as the predicate-argument structure is concerned, nonprototypical predication is treated in the same way as verbal clause constructions, more precisely as intransitive clause constructions (cf. Croft and Nivre, 2025). Different realizations of such constructions are further differentiated by the annotation of the predicate, which mainly depends on the type of construction, and by the annotation of copulas or other components of the strategy used to encode the predication. Let us now discuss each of these in turn.

In object predication, the predicate is a nominal expression denoting an object category,

Construction	UD Annotation
Object predication	PRED = NOUN
Property predication	PRED = ADJ
Locative predication	PRED = ADV
	PRED = NOUN[Case=Obl]
	PRED = NOUN $\xrightarrow{\text{case}}$ ADP

Strategy	UD Annotation
Zero	RE $\xleftarrow{\text{nsubj}}$ PRED
Nonverbal copula	RE $\xleftarrow{\text{nsubj}}$ PRED $\xrightarrow{\text{cop}}$ PRON
Verbal copula	RE $\xleftarrow{\text{nsubj}}$ PRED $\xrightarrow{\text{cop}}$ AUX
Verbal	RE $\xleftarrow{\text{nsubj}}$ VERB (?)

Table 1: UD annotation of different strategies for nonprototypical predication; RE = referring expression; PRED = predicate expression.

which means that the head word typically bears the part-of-speech tag NOUN and may have dependents such as determiners and nominal modifiers. In property predication, by contrast, the head word of the predicate is normally tagged ADJ and may take admodifers and other dependents typical of adjectives. In locative predication, finally, the predicate denotes a location, which means that it is either a locative adverb, tagged ADV, or an oblique nominal expression. The annotation of predicates in different constructions is summarized schematically in the top half of Table 1.

Turning next to the annotation of different strategies for encoding the predication, we first note that the zero strategy has no explicit annotation over and above the annotation of the predicate and the subject relation. This is illustrated in Figure 1(a), which shows the UD annotation of example (5a) from Tiwi.[2] For the two copula strategies, there is in addition a *cop* relation from the predicate to the copula, as seen in Figure 1(b–c), which shows the UD annotation of example (3) from Awtuw and example (2) from Amele. The only difference between these two examples is that the nonverbal copula is tagged PRON, while the verbal copula is tagged AUX. For the verbal strategy, finally, there are currently no explicit guidelines in UD, but the general guidelines for part-of-speech tagging seem to imply that the predicate should be tagged VERB because of the verbal morphology. This would make the annotation the same as for intransitive clauses, as seen in Figure 1(d), which shows the likely UD annotation of example (1b) from Moru. This would also mean that the VERB tag in this case "overrides" the tag implied by the construction, as shown in the bottom half of Table 1, which summarizes the annotation of different strategies.

In conclusion, it is clear that UD can handle all common strategies for nonprototypical predication using the *nsubj* argument relation together with the optional *cop* relation. In this way, the annotation also captures the constructional similarity across different strategies, in contrast to frameworks that treat copula verbs as heads. It is not clear, however, whether the annotation of the verbal strategy can distinguish nonprototypical predication from the intransitive clause construction.

3. Nonpredicational Clauses

In terms of information packaging, predicational clauses are topic-comment constructions, constructions that add new information (the comment) about one or more existing discourse referents (the topic). In terms of the file metaphor (Heim, 1983; Stassen, 1997), the predicational information packaging adds content to an existing file for a referent. However, languages also contain clause constructions that are based on other types of information

[2] In this and all following examples, we simplify the UD representations by omitting (a) lemmas and (b) morphological features that are not relevant for the discussion. For reasons of space, we do not repeat glosses and translations in the examples.

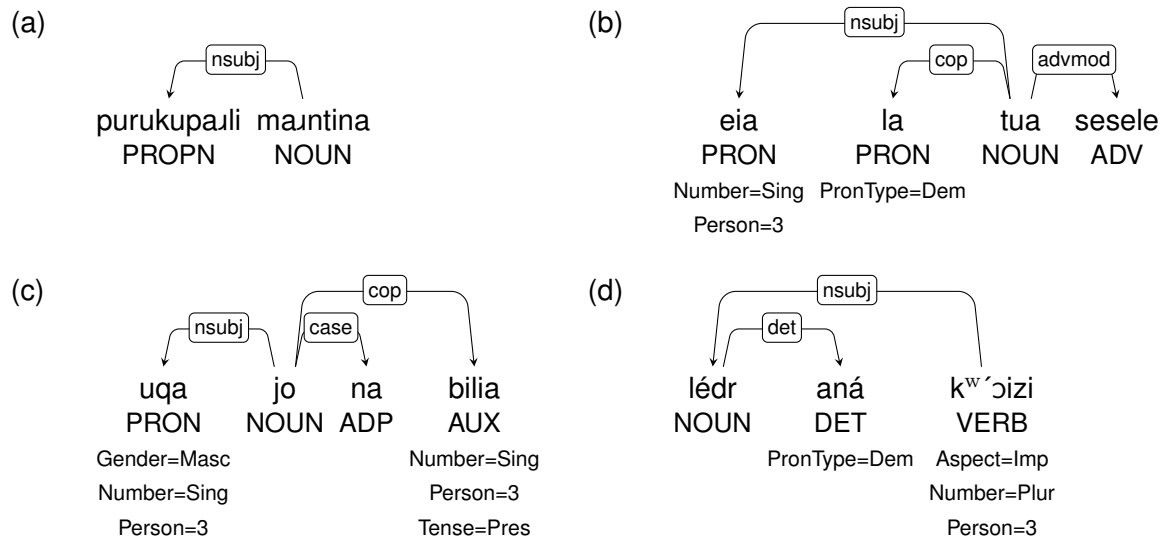

Figure 1: Simplified UD annotation of nonprototypical predication constructions/strategies: (a) object/zero, (b) object/nonverbal copula, (c) locative/verbal copula, (d) property/verbal.

packaging (Lambrecht, 1994).

In *thetic* clauses, there is no split between topic and comment, because all the information in the clause is new. A typical example is a presentational construction like *there's a cat*. In *identificational* clauses, two different characterizations are presented as pertaining to the same entity or event, as in the equational construction *the butler was the murderer*. In terms of the file metaphor, thetic constructions create a new file, while identificational constructions merge two existing files. We refer to thetic and identificational clauses collectively as nonpredicational clauses.

Nonpredicational constructions and strategies can be quite diverse and have not been studied as extensively as predicational clauses in linguistic typology. Our discussion will therefore be limited to the most important types discussed in Croft (2022). However, before we turn to these we will discuss a few topic-comment constructions that differ from standard predicational clause constructions.

3.1. Topic-Comment Constructions

In basic predicational clauses, the grammatical subject denotes the most topical (or salient) argument, but languages have a variety of mechanisms for promoting other arguments when required by the discourse context, including nonbasic voice constructions like passive-inverse and causative constructions, discussed in Croft and Nivre (2025). Another possibility is to use a *detached* phrase strategy, as in the following Swedish example:

(6) Lisa, henne har jag inte sett i dag
Lisa her have I not seen in day
'Lisa, I have not seen her today'

In this example, the detached topic phrase *Lisa* is preposed and co-referential with the object argument *henne*, but detached phrases can also be postposed, as in the following classic French example, where the detached phrase is co-referential with the subject (Maslova and Bernini, 2006):

(7) ils sont fous, ces romains
they are mad those Romans
'They're mad, those Romans'

A special kind of topic construction is the *hanging topic phrase*, where the topic does not correspond to an argument at all. This may be realized with an overtly coded strategy, as in example (8) from Japanese (Iwasaki, 2013), or with a zero-coded strategy, as in example (9) from Chinese (Li and Thompson, 1981).

(8) zoo wa karada ga okii
elephant TOP body NOM big:NPST
'the elephant – its body is big'

(9) xiàng bizi cháng
elephant nose long
'elephants' noses are long'

Hanging topics can be topics without being arguments because the sentence can be construed as about the referent of the hanging topic phrase, typically in virtue of a relationship like possession, exemplified in (8) and (9), or a figure-ground relation.

3.2. Thetic Constructions

The main characteristic of thetic constructions is that there is no division between topic and comment; all the information is new, including the referents. Among the clearest examples are *presentational* constructions, where the function is to open a discourse file for a referent, and possibly also to provide some information about this referent to the hearer. A typical case is the *presentational location* construction in (10), which introduces the cat as a new referent and contrasts with the *predicational location* construction in (11), where the existence of the cat is presupposed.[3]

(10) there is a cat on the mat
(11) the cat is on the mat

Presentational constructions are entity-central (Sasse, 1987), but there are also thetic constructions that are event-central, such as *weather* constructions, like *it's raining*.

The strategies for thetic constructions are varied but have in common that they make the phrase expressing the most topical referent look less like a subject and/or make the form expressing the event look less like a predicate (Croft, 2022, p. 339). It is very common to use *prosody* or *word order* to differentiate thetic constructions from otherwise similar predicational constructions, for example, by postposing the subject to the verb in SVO languages, as in *along comes a farmer with a donkey*.

Presentational constructions are often introduced by grammaticalized markers, as illustrated in the following examples:

(12) a. there is a cat on the mat
b. det sitt-er en katt på mattan
it sit-PRS a cat on mat-DEF.SG
'there is a cat on the mat'
c. hay un gato en el tapete
have.3SG.prs a cat on the mat
'there is a cat on the mat'

In the English example (12a), the locative expression *there* is a fixed marker of the presentational construction with the referring phrase postposed to the verb. In the Swedish example (12b), the third person neuter pronoun *det*, which has a variety of pleonastic uses in the language, is combined with the locational verb *sitta* (sit) to introduce the new referent. The Spanish example in (12c) represents another common pattern, namely the recruitment of a presentational possession construction to express presentational location, using the verb *hay*, an impersonal form of *haber* (have).[4]

A more complex strategy for thetic constructions is the *split* strategy (Sasse, 1987), where the encoding of the most prominent referent as less subject-like is combined with the predicate being in subordinate form. This is illustrated with a French example in (13) (Wehr, 1984; Sasse, 1987), where a typical presentational strategy is combined with a predicate expressed in the form of a relative clause.

(13) voilà la sirène qui hurle
there-is the siren which wails
'the SIREN is wailing'

Weather constructions, finally, exhibit a range of different strategies, three of which are exemplified below:

(14) a. llueve
rain.3SG.PRS
'it's raining'
b. es regnet
it rain:3SG.PRS
'it's raining'
c. akpɔkpɔ tin
cloud exist
'it's cloudy'

The Spanish example (14a) uses a verb that does not take any argument at all, while the German example (14b) uses a verb together with an impersonal, non-referential pronoun in subject position. The Gungbe example (14c),

[3] A similar contrast is found in possession clauses, where the *presentational possession* construction *she has a book* introduces a new referent of the book type, while the *predicational possession* construction *the book is hers* presupposes such a referent.

[4] Recruitment may also go in the opposite direction, so that presentational possession is expressed using a locative strategy. For more information on strategies for presentational possession, see Stassen (2009) and Croft (2022).

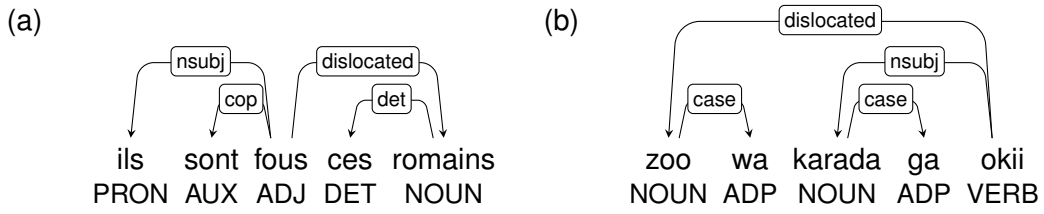

Figure 2: Simplified UD annotation of topic-comment constructions: (a) detached, (b) hanging topic phrase.

from Eriksen et al. (2012), instead uses a noun that combines with a general verb of existence.

3.3. Identificational Constructions

Identificational constructions divide the information in a clause in two, but instead of providing new information about a known topic, they equate two semantic components known from different descriptions. A typical case is the *equational* construction in (15), where the new information is not the existence of the butler or the murderer but the fact that the two are the same.

(15) the butler was the murderer

More generally, the two parts of an identificational construction are a *presupposed open proposition* (POP) with an unknown part X and a *focus* that provides the missing part X of the POP. In the equational construction, the focus is a refererring expression, but identificational constructions in general can involve any part of the information in a clause, as illustrated in the following examples from Croft (2022), where capitalization marks contrastive stress:

(16) I have ONE brother

(17) I HAVE finished my assignment

In (16), the POP is *I have X brothers* and the focus is *one* (not two or three). In (17), the unknown information is the polarity of the proposition, and the focus is a positive polarity.

The strategies for identificational constructions largely overlap with those used for other clause-level information packaging that differs from the default topic-comment packaging and again include prosody, exemplified above, and word order. A common overt morphosyntactic strategy is a *cleft* strategy, which divides the clause into two functional parts corresponding to the focus and the POP. Below are two Swedish examples of the cleft strategy:

(18) a. det var igår (som) hon kom
it was yesterday (that) she came
'it was yesterday (REL) she came'

b. vad som krävs är en ny start
what REL is-required is a new start
'what is required is a new start'

In (18a) the focus *igår* (yesterday) is linked by an equational copula to the POP, which takes the form of a relative clause even though it does not have the usual modifying function. In (18b), the POP is again expressed as a subordinate clause, but in this case it precedes the focus *en ny start* (a new start).

3.4. UD Annotation

The UD annotation framework for clause constructions focuses on the analysis of predicates, arguments and modifiers. This means that, even though core grammatical relations are defined partly in terms of information packaging, the framework lacks concepts for distinguishing different information packaging functions, such as thetic and identificational functions, unless they are realized through special morphosyntactic structures. In particular, anything using only prosody or word order cannot be represented in the UD annotation. In what follows, we will therefore limit ourselves to overtly coded strategies.

Starting with special topic-comment structures, the detached strategy and the hanging topic phrases are both annotated with the *dislocated* relation, as shown in Figure 2 for examples (7) and (8). The fact that the dislocated element is co-referential with an argument in (7) but not in (8) can be captured by subtyping the *dislocated* relation with an argument relation – *dislocated:nsubj* in this example. This is done only in a small minority of UD treebanks

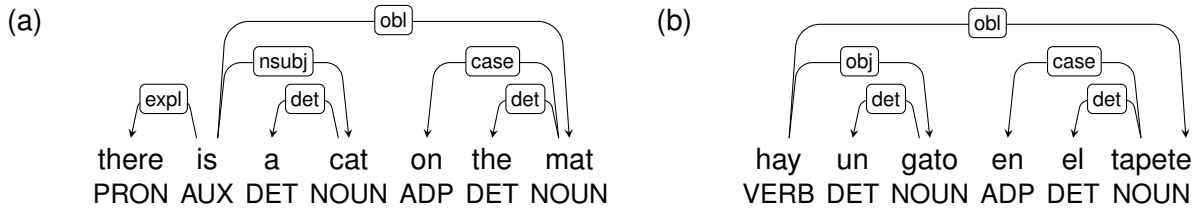

Figure 3: Simplified UD annotation of presentational location constructions.

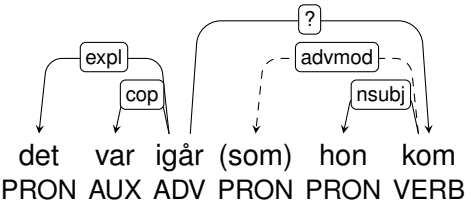

Figure 4: Simplified UD annotation of a cleft strategy.

but it could be recommmended more widely.

For thetic constructions, there are no general guidelines in UD, but the English presentational locational construction in (12a) is discussed together with nonverbal predication constructions. In this case, however, the guidelines say that the verb *be* should be treated as a main verb with existential meaning, not as a copula verb, as shown in Figure 3(a).[5] The referring phrase is treated as the nominal subject (*nsubj*) of this verb, and the grammaticalized marker *there* is attached with the *expl* relation, which is used in UD for expletive (non-referential) pronouns appearing in argument positions.[6] The same structure would be assigned to the Swedish example (12b), which is more expected given that the verb is the more contentful *sitta* (sit). The Spanish variant in (12c) is not explicitly discussed in the guidelines, but the current practice is to treat the referring phrase as the direct object (*obj*) of the impersonal verb *hay*, which is natural given the origin of the construction although it obscures the constructional similarity with the locational strategies used in English and Swedish. Given that the UD annotation prioritizes morphosyntactically encoded predicate-argument relations, divergences like these seem unavoidable, and fully capturing thetic constructions would most likely require either an additional annotation layer or a more complex subtyping system that can be used to annotate constructions and strategies in parallel. Such a system would also be needed to capture the split strategy, exemplified in the French example (13), which in the current system has to be analyzed as a regular relative clause construction.

Turning finally to identificational constructions, the situation is very similar in that there are no general guidelines for these. Equational constructions are essentially treated as non-prototypical predication, which means that *the butler was the murderer* is annotated exactly like *the butler was a doctor*. This is a reasonable analysis given the available tools, although it does not reflect the difference in information packaging. A notoriously difficult case is the cleft strategy, with all its variants, which has been discussed extensively in the UD community without a clear consensus emerging.[7] For the type of cleft exemplified by (18a), a common analysis is shown in Figure 4. This analysis treats the focus as the root of the clause, introduced by a copula verb with an expletive subject and followed by the POP in the form of a subordinate clause whose nature has been discussed extensively. Structurally, it

[5] https://universaldependencies.org/u/overview/simple-syntax.html#nonverbal-clauses

[6] The *expl* relation would also be used for the empty subject in the German weather construction in (14b).

[7] The official guidelines do not discuss clefts at all, but they are the subject of several ongoing discussions on the UD GitHub site.

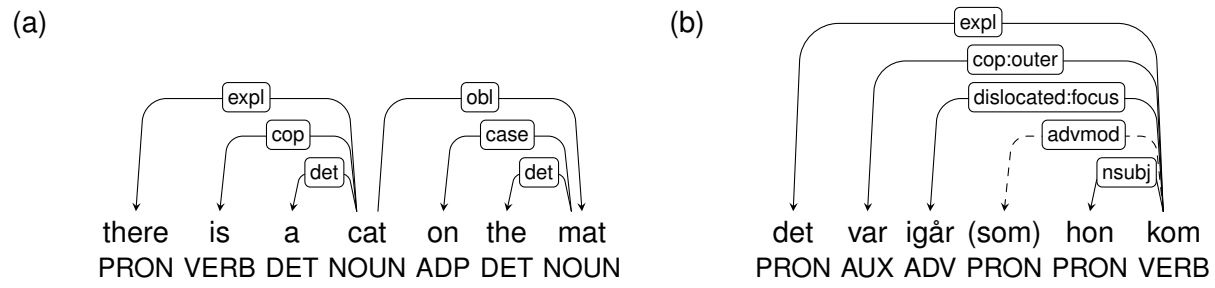

Figure 5: Revised annotation of (a) presentational and (b) identificational constructions.

is indistinguishable from a relative clause, but it is attached to a temporal adverb rather than a nominal, which raises the question whether it should be labeled *advcl* instead of *acl*. In addition, there is disagreement on whether it should be subtyped *:relcl*, to reflect the form of the clause, or *:cleft* to reflect its function.

4. Discussion

Our review of nonprototypical clause constructions has revealed a striking contrast between nonprototypical predication, which mostly has a straightforward analysis in UD (except possibly the verbal strategy, where some clarification is needed), and nonpredicational clauses, which are not discussed at all in the guidelines and where it is often unclear how to analyze specific constructions and strategies. We will therefore focus on the latter in our discussion.

What can be done to improve the analysis of thetic and identificational constructions, as well as non-prototypical topic-comment constructions? One approach is to add information on top of the UD annotation, as proposed by Weissweiler et al. (2024). Such annotations could, for example, add information about the focus part and the POP part of an identificational construction. A more radical approach would be to introduce new syntactic relations into the UD framework itself, like a *focus* relation for the analysis of identificational constructions. However, this would only be possible in a future version of UD, since the inventory of syntactic relations is fixed in UD v2.

Nevertheless, we believe that annotations could be improved even within the current version of UD by following the guiding principle that core argument relations should not be used to annotate strategies that primarily serve nonpredicational functions. This will sometimes require a change in what is considered the root of the construction, and it will often require an extended use of some (non-core) relations, possibly with special subtypes. For example, *dislocated* can be extended to cover all word order strategies, and *obl* can be generalized to all flagging strategies. In addition, special relations like *expl* and *cop* can be invoked for grammaticalized markers that do not function as real arguments or predicates. Let us consider two examples to illustrate the basic idea.

To recognize the thetic function of a presentational construction like (12a), one could treat the phrase introducing the new referent as the root of the construction and annotate the grammaticalized markers with *expl* and *cop*, as illustrated in Figure 5(a). This can be compared with the current analysis, shown in Figure 3(a), which (incorrectly) assumes that the referring phrase is a core argument of the copula verb. In cases where the presentational construction is instead realized with a lexical verb and a non-canonical word order, as in Spanish *viene Juan* (lit. “comes John”), a new relation subtype *dislocated:thetic* can be used to indicate the strategy (*dislocated*) as well as the construction type (*:thetic*).[8]

To capture the identificational function of the cleft strategy in (18a), one could introduce a new subtype *dislocated:focus* to relate the focus phrase to the POP and again use *expl* and *cop* for the initial grammaticalized markers, as illustrated in Figure 5(b). The internal struc-

[8] An alternative would be to use a subtype that captures the role of the displaced argument, although this would make the thetic construction indistinguishable from the special topic-comment structures discussed above.

ture of the POP is here analyzed analogously with the relative clause structure that has been recruited for construction, but the external relation makes clear that it does not have its normal modifying function. It can be discussed whether the root of the clause should be the POP, as in Figure 5(b), or the focus part, as in the old analysis in Figure 4, but treating the POP as the root has two advantages: it clarifies that the POP, despite its form, does not function as a subordinate clause, and it generalizes more naturally to equational constructions.

These are just two examples, and even their analysis is open to discussion, so a detailed proposal for guidelines for thetic and identificational constructions will have to be left for future work. However, we hope to have shown that simply annotating them as if they were predicational constructions may hinder rather than help the meaningful cross-linguistic comparisons that UD was created to support (de Marneffe et al., 2021).

5. Conclusion

In this paper, we have taken another step towards a constructicon for UD, in the sense of Nivre (2025), by reviewing the way UD annotates nonprototypical predication and nonpredicational clauses, following the taxonomy of Croft (2022), thereby extending the previous work on reference and modification (Nivre and Croft, 2025) and verbal predication (Croft and Nivre, 2025).[9] For nonprototypical predication, the guidelines provide a straightforward analysis of the main constructions and strategies, as outlined in Table 1. For nonpredicational clauses, the situation is much less clear, and many of the constructions and strategies discussed in Croft (2022) are either not covered at all in the current guidelines, or are subsumed under predicational constructions. This deserves further discussion, and we have briefly outlined a strategy for tackling the issue.

[9] In concurrent papers, we also review the annotation of complex predicates (Coneglian et al., 2026) and speech act constructions (Nivre et al., 2026).

6. Acknowledgments

This work received support from the CA21167 COST action UniDive, funded by COST (European Cooperation in Science and Technology), from the Swedish Research Council (grant no. 2022-02909), and from the Brazilian National Council for Scientific and Tecnhnological Development, CNPq (grant no. 153972/2025-4 and no. 405484/2025-9). We thank Elisabet Engdahl, Leonie Weissweiler and three anonymous reviewers for comments and suggestions that helped improve the article.

7. Bibliographic References

André V. Lopes Coneglian, Joakim Nivre, and William Croft. 2026. Complex predicates in Universal Dependencies. In *Proceedings of the Ninth Workshop on Universal Dependencies*.

William Croft. 2022. *Morphosyntax: Constructions of the World's Languages*. Cambridge University Press.

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.

Marie de Marneffe, Christopher D. Manning, Joakim Nivre, and Daniel Zeman. 2021. Universal Dependencies. *Computational Linguistics*, 47:255–308.

Pål Eriksen, Seppo Kittilä, and Leena Kolehmainen. 2012. Weather and language. *Language and Linguistics Compass*, 6:383–402.

Irene Heim. 1983. File change semantics and the familiarity theory of definiteness. In R. Bäuerle, C. Schwarze, and A. von Stechow, editors, *Meaning, Use, and Interpretation of Language*, pages 164–189.

Shoichi Iwasaki. 2013. *Japanese*. John Benjamins.

R. L. Johnston. 1980. *Nakanai of New Britain: The Grammar of an Oceanic Language*. The Australian National University.

Knud Lambrecht. 1994. *Information Structure and Sentence Form: Topic, Focus and the Mental Representations of Discourse Referents*. Cambridge University Press.

Charles Li and Sandra A. Thompson. 1981. *Mandarin Chinese: A Functional Reference Grammar*. University of California Press.

Arthur Lorenzi, Peter Ljunglöf, Ben Lyngfelt, Tiago Timponi Torrent, William Croft, Alexander Ziem, Nina Böbel, Linnéa Bäckström, Peter Uhrig, and Ely A. Matos. 2024. MoCCA: A model of comparative concepts for aligning constructicons. In *Proceedings of the 20th Joint ACL – ISO Workshop on Interoperable Semantic Annotation*, pages 93–98.

Elena Maslova and Giuliano Bernini. 2006. Sentence topics in the languages of Europe and beyond. In Giuliano Bernini and Marcia L. Schwartz, editors, *Pragmatic Organization of Discourse in the Languages of Europe*. Mouton de Gruyter.

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.

Joakim Nivre, André V. Lopes Coneglian, and William Croft. 2026. Speech act constructions in Universal Dependencies. In *Proceedings of the Ninth Workshop on Universal Dependencies*.

Joakim Nivre and William Croft. 2025. Reference and modification in Universal Dependencies. In *Proceedings of the 8th Workshop on Universal Dependencies (UDW))*, pages 1–10.

Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajič, Christopher D. Manning, Ryan McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty, and Dan Zeman. 2016. Universal Dependencies v1: A multilingual treebank collection. In *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC)*, pages 1659–1666.

Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Jan Hajič, Christopher D. Manning, Sampo Pyysalo, Sebastian Schuster, Francis Tyers, and Dan Zeman. 2020. Universal Dependencies v2: An evergrowing multilingual treebank collection. In *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC)*, pages 4034–4043.

C. R. Osborne. 1974. *The Tiwi Language*. Australian Institute of Aboriginal Studies.

John R. Roberts. 1987. *Amele*. Croom Helm.

Hans-Jürgen Sasse. 1987. The thetic-categorical distinction revisited. *Linguistics*, 25(5):11–80.

Leon Stassen. 1997. *Intransitive Predication*. Oxford University Press.

Leon Stassen. 2009. *Predicative Possession*. Oxford University Press.

Archibald N. Tucker. 1940. *The Eastern Sudanic Languages*. Oxford University Press.

Barbara Wehr. 1984. *Diskursstrategien im Romanischen*. Gunter Narr.

Leonie Weissweiler, Nina Böbel, Kirian Guiller, Santiago Herrera, Wesley Scivetti, Arthur Lorenzi, Nurit Melnik, Archna 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 2024)*, pages 16919–16932.