Complex Predicates in Universal Dependencies

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

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 the results of such a review focusing on complex predicates. We arrive at distinct findings regarding the two main types of complex predicates. The UD framework can well accommodate eventive complex predicates, particularly serial verbs, and more grammaticalized forms of complex predicates, such as voice and TAMP auxiliaries, with the exception of incorporating strategies. However, the guidelines for stative complex predicates could be revised based on the typology of morphosyntactic strategies. We briefly discuss possible ways in which UD can be extended to better capture these strategies.

Keywords: Universal Dependencies, eventive complex predicates, stative complex predicates

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 find out whether UD meets these requirements, Nivre (2025) proposes to build a constructicon for UD based on the survey of universal constructions and morphosyntactic realization 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 object predication construction is defined as any construction in any language where an object concept is predicated of a reference, as in the English *she is a doctor*. The verbal copula strategy describes an object predication construction that involves a linking element in the form of a verb, like English *is*. Other common strategies in the world's language are the nonverbal copula strategy (with a nonverbal linking element) and the zero strategy (with no linking element) (Croft, 2022; Nivre et al., 2026b).

The rationale behind the constructicon proposal is that, by checking if the current UD guidelines cover all attested strategies of a given construction, we can identify gaps in the guidelines and propose additions when needed. Moreover, in the same process, we can assess to what extent the annotations of different strategies capture the fact that they are realizations of the same construction, which facilitates cross-linguistic comparisons in the spirit of UD's "mixed structural-functional system" for annotation.[1] A further step would be to investigate how actual treebanks have dealt with the phenomena at hand. However, although we reference specific treebanks when relevant, a systematic review of treebanks lays outside the current scope.

Previous and concurrent contributions to this project include studies of reference and modification constructions (Nivre and Croft, 2025), verbal predication constructions (Croft and Nivre, 2025), nonprototypical predication and nonpredicational clauses (Nivre et al., 2026b), and speech act constructions (Nivre et al., 2026a). In this paper, we extend the review to complex predicates, a family of constructions discussed in Chapters 13–14 of Croft (2022).

Complex predicates are monoclausal constructions in which two or more elements com-

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

bine into a single predication, in contrast to single predicates in which predication is expressed by one element (Croft, 2022). They cover a wide range of semantic categories and present a diverse set of morphosyntactic structures. Complex predicates are intermediate structure types in the grammaticalization path, between biclausal structures and affixes (Bowern, 2008). Following Croft (2022), we consider two main groups of complex predicates, namely, *eventive complex predicates*, where both elements of the predicate denote processes (Section 2), and *stative complex predicates*, where the other element denotes a state (Section 4). We also consider the more grammaticalized types of eventive predicates, such as voice and TAMP auxiliaries (Section 3).

2. Eventive Complex Predicates

The main characteristic of eventive complex predicates is that they denote one single event. Therefore, the elements of a complex predicate share the same argument structure, and the same semantic values with respect to tense, aspect, evidentiality, mood, modality and reality status. Additionally, a complex predicate constitutes a single assertion, thus the complex predicate will be negated or questioned as a whole (Croft, 2022). The wide range of constructions and morphosyntactic strategies for eventive complex predicates covers different areas of the grammar of languages, such as serial verbs, light verbs and support verbs, and verb-object idiomatic expressions. Here, we focus exclusively on *serial verb* strategies, that is, strategies that combine one or more verbs, each denoting an event, to realize eventive complex predicates expressing meanings such as direction, orientation, aspect, change of state, sequence of subactions, simultaneity of actions, among others (Aikhenvald, 2018, p. 239).

2.1. Strategies for Serial Verbs

The range of morphosyntactic strategies for serial verbs can be defined along three morphosyntactic parameters (Croft, 2022): (a) contiguity—whether or not the predicate elements are contiguous; (b) boundedness or incorporation—whether or not the predicate elements form one single word; (c) locus of predicate inflection—which element(s) receives predicate inflection. Starting with contiguity, example (1), from Tetun Dili (Hajek, 2006; Aikhenvald, 2018, p. 2), illustrates a case of non-contiguous serial verbs, in which they are separated by an argument phrase; while example (2), from Barai (Aikhenvald, 2018), illustrates a case of serial verb contiguity.

(1) tuda bola mai
throw ball **come**
'throw the ball over here'

(2) fu fase naebe fi isoe
he letter NEG **sit write**
'he did not sit and write a letter'

In examples (1–2) each element of the serial verb is a separate word. By contrast, example (3), from Tepehua (Watters, 2017; Aikhenvald, 2018, p. 98) illustrates a case of incorportion, where the two elements of the serial verb form a single word. Example (4), from Sakao (Guy, 1974; Durie, 1997, p. 303), illustrates a case in which the serial verbs are bound, but an argument is interpolated between the two forms in boldface.

(3) haɬtanah-tsuku-ɬ=ta
walk-**begin**-PFV=already
'S/he began to walk' [lit. walk begin]

(4) ɣa-βǿt-nœð-p-ri-lam l-aðaɣ
3PL-**pluck**-coconut-PFV-DIR-**come** to-baskets
'They plucked coconuts hither into the baskets'

With respect to the locus of predicate inflection, there are five morphosyntactic strategies (Durie, 1997; Croft, 2022): inflection on the first verb; inflection on the last verb; same inflection on all verbs; split inflection; separate inflection. All these types have been attested cross-linguistically. Due to space limitations, we cannot provide examples for all these strategies, but the Òbòlò example (5) illustrates the same inflection on both verbs (Durie, 1997, p. 299).

(5) é-gwên èmì é-nû
PL-**call** 1SG PL-**come**
'let them call me to come'

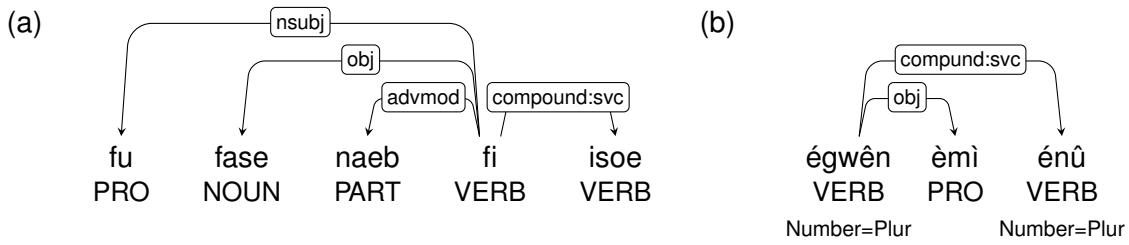

Figure 1: Simplified UD annotation of serial verb strategies: (a) contiguous serial verb strategy, (b) non-contiguous serial verb strategy with the same inflection on both verbs.

The strategies discussed up to this point are all zero-coded serial verb strategies, that is, there is no element that links the predicate elements. However, there are cases of overt coding for a serial verb strategy, by means of a conjunction element or a converb construction. The following example from Mooré (Schiller, 1990, p. 38) is a case of a serial verb that optionally takes a conjunction.

(6) a ị̀ku sụ̀ugā (n) wāg nemdā
he took knife (CNJ) cut meat
'he cut the meat with a knife'

Even though converb constructions typically express a semantic relation between two events in biclausal structures, Bisang (1995) and Shibatani (2009) show they may also be used to express a semantic relation between events that form one single event, as is the case of serial verb strategies. The Khalkha Mongolian example (7), from Street (1963), cited in Bisang (1995, p. 170), shows a converb construction in which the converb is used as a flag, and the verb *give* for the R role of the communication event (Croft, 2022, p. 410).

(7) tüünig duuda-ž ög-öörej
him phone-**CNV** **give**-IMP
'please call him [to the phone] for [me].'

Overt linking in serial verbs requires careful analysis, because conjunction and converb constructions can also be found in biclausal and multi-eventive predicates (Croft, 2022, ch. 15–19).

2.2. UD Annotation

According to the UD guidelines, serial verbs should be annotated with the relation *compound:svc*, that is, as a *compound* with the subtype *svc*, as Figure 1 illustrates.[2] In both these examples, the first element of the serial verb is treated as the head, but the guidelines allow any of the two elements to be treated as the head, based on language-specific criteria, and head-final serial verbs are found, for example, in UD treebanks for Turkic languages (Akhundjanova et al., 2025).

There are no explicit guidelines for the morphological annotation of serial verb elements, but it seems natural to assume that they are all assigned the part-of-speech tag VERB and morphological features in accordance with the inflection strategy, as illustrated in Figure 1(b). Finally, it should be stated that serial verb strategies involving incorporation pose a significant challenge for the current UD annotation, particularly bounded serial verb strategies, as in example (3), because the verbal elements of the serial verb tend to constitute only one word. We return to this issue in Section 5.

3. Grammaticalized Constructions

As mentioned above, eventive complex predicates may grammaticalize into constructions with different functions. Here we discuss two commonly attested cases, namely nonbasic voice constructions and TAMP-auxiliary constructions.

3.1. Nonbasic Voice Constructions

Eventive complex predicates may be used to express nonbasic voice, such as passive-

[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.

inverse, causative and applicative (Croft, 2022; Croft and Nivre, 2025). Typically, verbs like 'give' and 'take' grammaticalize into oblique flags, coding the relationship between argument phrases in the overall predication. This is the case in many languages around the world. Example (8), from Mandarin, illustrates the form *gěi* 'give' expressing the beneficiary of the action 'make' (Li and Thompson, 1985; Croft, 2022).

(8) māma gěi wǒ zuò jiaǒzi
mother **for** 1SG make dumpling
'mother made dumplings for me'

One problem that arises in the analysis of these constructions is the categorization of forms like *gěi* in Mandarin and in other languages, given that these forms are undergoing grammaticalization but may retain semantic content and morphosyntactic structure from their source. Due to the gradualness of grammaticalization, it may therefore be unclear whether forms like *gěi* should be analyzed as forming a complex predicate with a(nother) verb or as a flag combining with an argument phrase (Croft, 2022).

3.2. TAMP-Auxiliary Constructions

Auxiliaries are another common type of grammaticalization of eventive complex predicates, which come to express meanings such as tense, aspect, mood, modality and polarity (Givón, 2001; Anderson, 2006). They do not constitute a basic type of eventive complex predicate; however, they do share the same set of parameters and morphosyntactic strategies that were discussed for serial verbs in Section 2. Thus, auxiliary constructions, too, are defined by the parameters of contiguity, boundedness and locus of predicate inflection.

The English example *she might come* illustrates contiguity of auxiliary and main verb. By contrast, the auxiliary in the Kisi example (9) is non-contiguous with the main verb.

(9) fàlà có lɛ́ɛ́ŋdó yìkpàá
Fala **AUX** machete **sharpen**
'Fala is sharpening the machete'

Concerning incorporation, auxiliaries may be bound or free. The auxiliary in (9) is free, while the Abkhaz auxiliary in (10) is bound (Hewitt, 1979; Bybee et al., 1994).

(10) s-cà-r-o-w+p'
1SG-**go**-**COND**-be-STAT
'I must go'

Concerning locus of inflection, auxiliaries can be classified into the same five types as serial verbs (Croft, 2022, p. 415–416): (i) inflection on the auxiliary; (ii) inflection on the main verb; (iii) same inflection on auxiliary and main verb; (iv) split inflection; (v) separate inflection.

3.3. UD Annotation

The annotation of argument structure and voice constructions that result from the grammaticalization of eventive complex predicates requires language-specific parameters for guidelines. Consider example (8) above. From our constructional perspective, it is not clear which part-of-speech tag should be assigned to the form *gěi*, due to the fact that it is undergoing grammatical change. However, if the preferred option is ADP, then the syntactic relation that links it to the pronoun form has to be *case*. This is the analysis adopted in a current UD treebank for Chinese, (Poiret et al., 2023) and the corresponding UD annotation of (8) is given in Figure 2(a).

The annotation of auxiliary constructions is straightforward in UD. Basically, all the cases mentioned in the preceding discussion are tagged as AUX and linked to the main verb with *aux* relations, which is reserved for auxiliaries of tense, aspect, evidentiality, or mood to the extent that they are grammaticalized in the language (De Marneffe et al., 2024). In UD v2, it is recommended that the subtype *aux:pass* be used for grammaticalized periphrastic voice, as in the Brazilian Portuguese example (11), illustrated in Figure 2(b), where the verb form *foi* is the passive voice auxiliary.

(11) o vaso foi quebrad-o
DEF.M.SG vase be.3SG.PRF broken-M
'the vase was broken'

Although the guidelines for auxiliaries are clear in themselves, the gradual nature of grammaticalization may give rise to borderline cases, and less grammaticalized auxiliary strategies are generally not to be analyzed with the AUX tag and *aux* relation. A typical case is what has been called periphrastic aspect, as in *I*

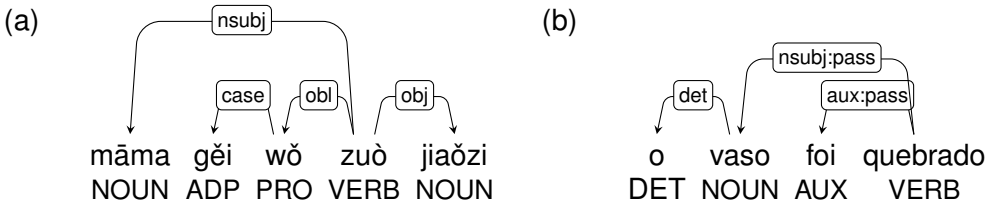

Figure 2: Simplified UD annotation of grammaticalized constructions: (a) nonbasic voice, (b) auxiliary.

began to wonder. In cases like this, the recommendation is that the verb *began* be analyzed as a main verb taking a complement clause headed by the verb *wonder*; the relation from *began* to *wonder* is labeled *xcomp*, because there is obligatory control, and *to* is linked to *wonder* with the *mark* relation, which indicates syntactic subordination.

4. Stative Complex Predicates

A stative complex predicate is a monoclausal construction in which one element denotes a process and the other denotes a state (Croft, 2022). The stative concept does not modify the referent, as is the case in attributive modifying constructions (Nivre and Croft, 2025), nor does it constitute the primary element of the predication. Rather, the stative element constitutes a part of the complex predicate. The three main semantic types of stative complex predicates are: *resultative, depictive*, and *manner*. The English examples (12-14) from Croft (2022, p. 432) illustrate these three major types, respectively.

(12) we **painted** the wall **red**

(13) I **ate** the carrots **raw**

(14) we **crawled** down the slope **slowly**

Most of the typological and cross-linguistic work on stative complex predicates is focused on depictive and manner predicates, while resultative predicates still lack robust cross-linguistic investigation.

4.1. Strategies for Stative Complex Predicates

The three types of stative complex predicates share many morphosyntactic strategies (Schultze-Berndt and Himmelmann, 2004; Loeb-Diehl, 2005; Croft, 2022), which can be characterized according to three morphosyntactic parameters (Loeb-Diehl, 2005; Croft, 2022): (i) whether it uses balanced or deranked[3] predicates; (ii) whether the stative predicate is indexed or non-indexed; and (iii) whether the stative predicate recruits[4] the action predication construction – what Loeb-Diehl (2005) calls "verby" – or the adjective or nominal construction – called "nouny".

Let us start with the four verby strategies, i.e. strategies that recruit the action predication construction (Croft, 2022). Two of these, the *coordinate personal strategy* and the *coordinate impersonal strategy*, are balanced strategies, where the stative predicate constitutes a separate coordinated predication. The difference between the two is that, in the former the stative predicate indexes the participant (hence, personal), while in the latter the stative predicate does not (hence, impersonal). Example (15) from Muna (van den Berg, 1989; Loeb-Diehl, 2005) and example (16)[5] from Koasati (Kimball, 1985; Loeb-Diehl, 2005) illustrate the personal and impersonal coordinate strategies, respectively.

(15) ne-rimba no-tende
3SG.RL-be_fast 3SG.RL-run
'he runs fast' [lit. 'he is fast (and) he runs']

[3] In complex predicates, a balanced strategy recruits the construction of a simple predicate construction with inflections, etc. By contrast, a deranked strategy either lacks the inflections of the predicate, uses a different set of inflections from the predicate, or has designated affixes that overtly code the other predicate (Croft, 2022).

[4] Recruitment is a relationship between two constructions in which the structure of one is recruited or whose function is extended to use in another construction (Croft, 2022).

[5] In example (16), DS stands for 'different subject' (Loeb-Diehl, 2005).

(16) ca-conoská-k pálk-á:hi:si-n bókl
1SG.POSS-heart-SBJ **be_fast-ADV-DS** beat
'my heart is beating very fast'
[lit. 'my heart is fast and beats']

By contrast, *participial* and *converb strategies* are deranked strategies. The difference between them is indexation: while the first indexes a participant of the main event, the latter does not. In example (17), from Sanuma (Borgman, 1990; Loeb-Diehl, 2005), manner is expressed by a same-subject participle marked by *-i*, glossed as REL. Example (18), from Turkana (Dimmendaal, 1983; Loeb-Diehl, 2005), illustrates the converb strategy, where *nɪ* indexes the event, not the participant.

(17) opi-i a kali-palo-ma
be_slow-**REL** 3SG work-REPET-COMPLET
'he worked slowly'

(18) è-pès-e-tè nɪ-a-ron-o-ɲì
3-kick-A-PL **REL.N-INF-be_bad-SG.REL**
'they kick him badly'
[lit. they kick him, it being bad]

We now turn to the four nouny strategies, i.e. the strategies that recruit adjectival or nominal constructions and use inflectional categories like gender, number, and case. The *adjective personal strategy* recruits a modifying phrase construction (Nivre and Croft, 2025). This strategy indexes an event participant, not the event itself (Croft, 2022). Loeb-Diehl (2005) notes that this strategy is rather uncommon cross-linguistically. It is observed in Classical Latin and some daughter languages, such as Spanish, Italian, Romanian, and Portuguese. In the Spanish example (19), the adjective forms *tranquilos* and *felices* index the plural (Slager, 1995; Loeb-Diehl, 2005).

(19) aqui vivimos tranquilos y felices
here live.1PL quiet.**PL** and happy.**PL**
'we live quietly and happily here.'
[lit. 'we live here quiet and happy']

The *adjective impersonal strategy* differs from the former only with respect to indexation. This strategy tends to be used for manner predicates, and not for depictives; as Loeb-Diehl (2005) notes, this strategy is more event-oriented. This strategy is also found in Romance languages, like Portuguese, among many other languages, illustrated in (20). In this example no indexation is observed in the adjective form *rápido* (fast).

(20) ela corre rápido
3SG.F run.3SG.PRS **fast**
'she runs fast/fastly'

In deranked strategies the stative predicate may combine with a copula or be expressed with a flag. In the first case, the *copular participle strategy*, the copula is a deranked form combined with an adjectival form. In the Malayalam example (21) the bound *-aayi* is a past participle copular form (Askher and Kumari, 1997; Loeb-Diehl, 2005). In the second case, the *adpositional strategy*, the stative predicate is expressed with a flag like a referring phrase. In the Japanese example (22), the adjective form is marked with a locative and instrumental flag (Takezawa, 1993; Schultze-Berndt and Himmelmann, 2004).

(21) aval bhamgiy-aayi prasamgiccu
she beauty-**PTCP.COP** speak-PST
'she spoke beautifully'

(22) John-ga sakan-o hadaka-de tabe-ta
John-NOM fish-ACC nude-**INS** eat-PST
'John ate the fish nude'

The *adpositional personal strategy* (Schultze-Berndt and Himmelmann, 2004) involves an indexed stative predicate, similarly to the other personal strategies discussed above. The Russian example (23) illustrates the use of both a flag marker and a personal marker in the stative predicate element (Nichols, 1978; Schultze-Berndt and Himmelmann, 2004).

(23) on umer molodym
he die.PST.**M.SG** young.**M.SG**.INS
'he died young'

The *predicational strategy* is one in which the manner or stative item becomes the main predicate. As Loeb-Diehl (2005) describes it, in this strategy the dynamic subevent is nominalized and is the subject of the sentence. This strategy is very rare crosslinguistically and is illustrated in example (24) from Mokilese (Harrison, 1976; Loeb-Diehl, 2005).

(24) ah kijou dahr
his run fast
'he runs fast' [lit. 'his running is fast']

Finally, the *adverbial strategy* uses a distinct morphosyntactic form to express manner complex predicates. It typically involves an affix

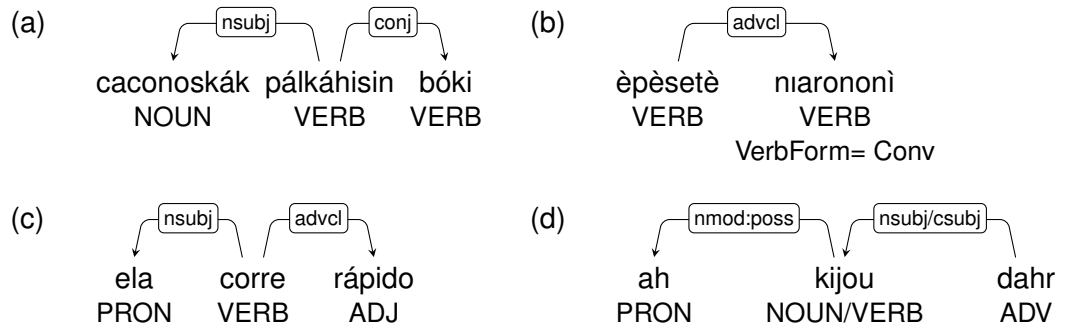

Figure 3: Simplified UD annotation of stative complex predicate strategies: (a) coordinate strategy, (b) converb strategy, (c) adjective strategy, (d) predicational strategy.

to the manner predicate, like the *-ly* suffix in English, or *-mente* in Portuguese. There is a specially designated word form for this strategy (Loeb-Diehl, 2005; Croft, 2022).

4.2. UD Annotation

The current UD guidelines do not treat stative complex predicates as a concept, which means that different strategies are annotated based on their morphosyntactic form and are therefore often assimilated to other constructions that use similar strategies. Thus, in coordinate strategies, the stative predicate is tagged VERB and linked to the main predicate with the *conj* relation, which is used to annotate all types of coordination in UD. This is illustrated in Figure 3(a), which shows the UD annotation of example (16).

In the participial, converb and adjective strategies, the stative predicate is tagged VERB or ADJ and linked to the main predicate with an *advcl* relation – a relation that is also used for full adverbial clauses and therefore makes the complex predicate constructions look like multi-clausal constructions. This is illustrated in Figure 3(b–c), which shows the UD annotation of examples (18) and (20).

In the adpositional strategy, the stative predicate is tagged ADJ (or possibly NOUN) and linked to the main predicate with an *obl* relation, and the flag is represented either by a Case feature, in case it is inflectional, or with a *case* dependent, in case it is an independent word. This groups this strategy together with non-core arguments of predicates.

In the adverbial strategy, the stative predicate is tagged ADV and linked to the main predicate with an *advmod* relation, which is not a clausal relation and therefore compatible with the monoclausal nature of the complex predicate construction, although it is also used for other types of construction, notably negation in many languages.

For the predicational strategy, finally, it is clear that the nominalized main predicate has to be analyzed as a subject of the stative predicate, but depending on whether the former is tagged NOUN or VERB, the subject relation may be either *nsubj* or *csubj*.

5. Discussion

The review carried out in this paper can be summarized in two main points. First, from a typological perspective, complex predicate constructions exhibit a wide range of possible annotations within the UD framework. This is explained, at least in part, by the fact that complex predicates are very broadly defined in the constructional framework of Croft (2022). Secondly, regarding the current UD guidelines, we found that UD v2 presents clear instructions for the annotation of most of the strategies covered in this paper, to varying degrees of comprehensiveness and typological adequacy. The current version of the guidelines for eventive complex predicates, particularly those more grammaticalized as TAMP auxiliaries, are the most consistent with typological findings. The guidelines for annotating stative complex predicates are more heterogeneous in that different strategies are grouped with different construction types. In particular, there is a clear divide between strategies that are

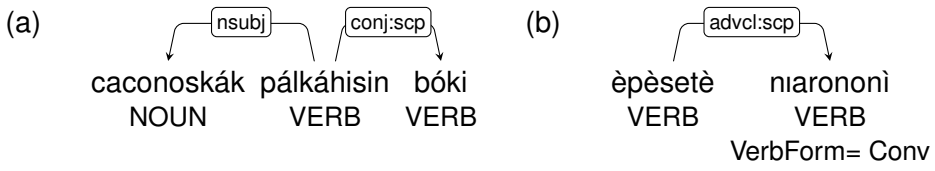

Figure 4: Revised UD annotation of stative complex predicate strategies with subtypes: (a) coordinate strategy, (b) converb strategy.

annotated as mono-clausal versus those that are analyzed as bi-clausal.

The prescription for the annotation of serial verb constructions is very clear in UD guidelines. However, as we noted above, the current version of the guidelines does not cover incorporating strategies, in which the elements form a single word, as is typical in more polysynthetic languages. This means that, at this time, we cannot present any conclusive proposal regarding incorporating strategies for serial verbs and for auxiliaries. The UD framework still lacks clear orientations for how to deal with incorporation and polysynthesis in general, despite attempts to establish guidelines (Tyers and Mishchenkova, 2020; Park et al., 2021; Koshevoy et al., 2023). Similarly, current guidelines cover only zero coded serial verb strategies, and give no indication for the annotation of the overt coding strategies. Since in these cases the overt element is typically a conjunction or linking element, we see no problem in tagging this element as CCONJ,[6] but it is not clear whether it can be assigned a *cc* relation without also changing the relation between the two predicates to *conj*. The latter would assimilate these strategies with the coordinate strategies for stative complex predicates, but would obscure the relation to zero coded serial verb strategies.

With respect to stative complex predicates, we have seen that the current guidelines can handle all the different strategies, but the relations that have to be used group them together with quite different constructions. Coordinate strategies are grouped with other types of coordination; participial, converb and adjective strategies are grouped with adverbial clauses; and adpositional strategies are grouped with oblique arguments. One way of mitigating this problem would be to introduce a new subtype *:scp* (for stative complex predicate), which would at the same time capture their constructional similarity and differentiate them from other constructions annotated with the same relations. This would mean using *conj:scp* for coordinate strategies, *advcl:scp* for participial, converb and adjective strategies, and *obl:scp* for the adpositional strategy. The former two cases are illustrated in Figure 4.

It is also worth noting that the use of the *advcl* relation for several strategies implies a bi-clausal structure, which appears incompatible with the definition of complex predicates as mono-clausal. It is particularly striking that the adverbial and the adjective strategy for manner predicates end up getting very different analyses, despite being very similar in nature. One way of mitigating this would be to allow the adjective strategy to be analyzed using the *advmod* relation, at least when used to express a manner predicate. However, it seems more difficult to avoid the use of *advcl* for participial and converb strategies, where the stative predicate is verbal. Hence, a fully satisfactory solution to this problem would probably require adding new universal relations, which can only be considered for future versions of UD.

6. 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 complex predicates, following the taxonomy of Croft (2022), thereby extending the previous work on reference and modification (Nivre and Croft, 2025) and verbal predication

[6] Our survey based on Croft (2022) has shown that only balanced strategies are used in this case. In any case, if there are languages in which overt coding is used, then the annotation should be different.

(Croft and Nivre, 2025).[7] On the one hand, for both serial verbs and auxiliaries, the current guidelines provide very straightforward instructions of annotation, with the limitation that they do not have any clear way of dealing with incorporation; however, this is due to the fact that the UD framework does not accommodate well phenomena related to polysynthesis. On the other hand, the guidelines for stative complex predicates are less satisfactory in that different strategies end up being grouped with different constructions and in many cases lead to a bi-clausal analysis. We have discussed a few ways in which these problems can be mitigated, but a complete revision of the guidelines will likely only be possible when a new major version of UD is launched.

7. 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 Leonie Weissweiler and three anonymous reviewers for comments and suggestions that helped improve the article.

8. Bibliographic References

Alexandra Aikhenvald. 2018. *Serial verbs*. Oxford University Press.

Arofat Akhundjanova, Furkan Akkurt, Bermet Chontaeva, Soudabeh Eslami, and Cagri Coltekin. 2025. Parallel Universal Dependencies treebanks for Turkic languages. In *Proceedings of the Eighth Workshop on Universal Dependencies (UDW, SyntaxFest 2025)*, pages 129–136, Ljubljana, Slovenia. Association for Computational Linguistics.

Gregory DS Anderson. 2006. *Auxiliary Verb Constructions*. Oxford University Press.

R. E Askher and T.C. Kumari. 1997. *Malayalam*. Routledge.

Walter Bisang. 1995. Verb serialization and converb – differences and similarities. In Ekkehard König and Martin Haspelmath, editors, *Converbs in Cross-Linguistic Perspective*, pages 137–88. Mouton de Gruyter.

Donald M. Borgman. 1990. Sanuma. In *Handbook of Amazonian Languages*, pages 15–248. Mouton de Gruyter.

Claire Bowern. 2008. The diachrony of complex predicates. *Diachronica*, 25(2):161–185.

Joan Bybee, Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. University of Chicago Press.

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.

Marie-Catherine De Marneffe, Joakim Nivre, and Daniel Zeman. 2024. Function words in universal dependencies. *Linguistic Analysis*, 43(3–4):549–588.

Gerrit Jan Dimmendaal. 1983. *The Turkana Language*. Foris.

Mark Durie. 1997. Grammatical structures in verb serialization. In Alex Alsina, Joan Bresnan, and Peter Sells, editors, *Complex Predicates*, pages 289–354. CSLI.

Talmy Givón. 2001. *Syntax, vol. 1*. John Benjamins.

J. B. M. Guy. 1974. *A Grammar of the Northern Dialect of Sakao*. Research School of Pacific Studies, Australian National Univeristy.

[7] In concurrent papers, we also review nonprototypical predication and nonpredicational clauses (Nivre et al., 2026b), as well as speech act constructions (Nivre et al., 2026a).

John Hajek. 2006. Serial verbs in Tetun Dili. In *Serial Verb Constructions: A Cross-Linguistic Typology*, pages 239–253. Oxford University Press.

S. P. Harrison. 1976. *Mokilese Reference Grammar*. University Press of Hawaii.

B. G Hewitt. 1979. *Abkhaz*. North-Holland.

Geoffrey D. Kimball. 1985. *Koasati Grammar*. University of Nebraska Press.

Alexey Koshevoy, Anastasia Panova, and Ilya Makarchuk. 2023. Building a universal dependencies treebank for a polysynthetic language: the case of Abaza. In *Proceedings of the Sixth Workshop on Universal Dependencies (UDW, GURT/SyntaxFest 2023)*, pages 1–6.

Charles Li and Sandra Thompson. 1985. *Mandarin Chinese: A Functional Reference Grammar*. University of California Press.

Flora Loeb-Diehl. 2005. *The Typology of Manner Expression*. Ph.D. thesis, Radboud University Nijmegen.

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.

Johanna Nichols. 1978. Secondary predicates. In *Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society*.

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. 2026a. 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, William Croft, and André V. Lopes Coneglian. 2026b. Nonprototypical predication and nonpredicational clauses in Universal Dependencies. In *Proceedings of the Ninth Workshop on Universal Dependencies*.

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.

Hyunji Hayley Park, Lane Schwartz, and Francis Tyers. 2021. Expanding universal dependencies for polysynthetic languages: A case of St. Lawrence Island Yupik. In *Proceedings of the First Workshop on Natural Language Processing for Indigenous Languages of the Americas*, pages 131–142.

Rafaël Poiret, Tak-Sum Wong, John Lee, Kim Gerdes, and Herman Leung. 2023. Universal dependencies for Mandarin Chinese. *Language Resources and Evaluation*, 57(2):673–710.

Eric Schiller. 1990. *An autolexical account of subordinating serial constructions*. Ph.D. thesis, Department of Linguistics, University of Chicago.

Eva Schultze-Berndt and Nikolaus P Himmelmann. 2004. Depictive secondary predicates in crosslinguistic perspective. *Linguistic Typology*, 8(1).

Masayoshi Shibatani. 2009. On the form of complex predicates: toward demystifying serial verbs. In Johanes Helmbrecht, Yoko Nishina, Yong-Min Shin, Stavros Skopeteas, and Elisabeth Verhoeven, editors, *Form and Function in Language Research*, pages 255–282. Mouton de Gruyter.

E. Slager. 1995. *Grammatica Spaans*. Spectrum.

John C. Street. 1963. *Khalkha Structure*. Indiana University Press.

Koichi Takezawa. 1993. Secondary predication and locative/goal phrases. In *Japanese Syntax in Comparative Grammar*. Kurosio.

Francis Tyers and Karina Mishchenkova. 2020. Dependency annotation of noun incorporation in polysynthetic languages. In *Proceedings of the Fourth Workshop on Universal Dependencies (UDW 2020)*, pages 195–204.

René van den Berg. 1989. *A Grammar of the Muna Language*. Foris.

James K Watters. 2017. Verb-verb compounds and argument structure in Tepehua. In *Argument Realisation in Complex Predicates and Complex Events*, pages 277–304. John Benjamins Publishing Company.