Elliptical Constructions in Estonian UD Treebank

Kadri Muischnek
University of Tartu
Institute of Computer Science/
Institute of Estonian and General Linguistics
Kadri.Muischnek@ut.ee

Liisi Torga
University of Tartu
Institute of Computer Science
Liisi.Torga@ut.ee

Abstract

This contribution is about the annotation of sentences with verbal predicate ellipsis in the Estonian Universal Dependencies (UD) treebank. The main aim of the UD initiative is to develop a cross-linguistically consistent treebank annotation scheme and build a multilingual treebank collection. There are more than 70 treebanks in over 100 languages in UD treebank collection version 2.2. However, the UD annotation scheme is constantly improved and amended and so the annotation of the treebanks is also changing from version to version.

Our article studies a problematic issue in representing syntactic structure – clauses with predicate verb ellipsis. UD syntactic annotation scheme is based on dependency syntax and as the dependency structure is verb-centered, predicate verb ellipsis causes more annotation problems than other types of ellipsis. We focus on such constructions (referred to in English often as gapping and stripping) in Estonian and their annotation in Estonian UD treebank versions 1 and 2.2.

Kokkuvõte

Artikkel käsitleb elliptilise öeldisega laustete märgendamist eesti keele *Universal Dependencies'* (UD) puudepangas. UD eesmärgiks on esiteks töötada välja puudepankade morfoloogilise ja sõltuvussüntaktilise märgendamise skeem, mis oleks keelest sõltumatu selles mõttes, et sobiks kõigi keelte märgendamiseks, ja teiseks luua selle märgendusskeemi järgi annoteeritud puudepankade kollektsioon. UD versioon 2.2 sisaldab enam kui 100 puudepanka rohkem kui 70 keeles. Märgendusskeemi arendatakse ja täiustatakse pidevalt ja seega tuleb UD kujul olevaid puudepanku uute versioonide tarbeks pidevalt ümber märgendada.

UD süntaktiline märgendus põhineb sõltuvussüntaksi põhimõtetel, mille järgi lause keskmeks on öeldis, tavaliselt finiitne verbivorm. Erandiks on koopulalaused (eesti keele puhul *olema*-verbiga laused), mille kõrgeimaks ülemuseks on mitte-verbiline predikaat. Kuna UD süntaksimärgendus on oma olemuselt verbikeskne, põhjustab öeldise ellips lausepuu moodustamisel rohkem probleeme kui mõne muu lausemoodustaja väljajätt. Tavaline on öeldisverbi väljajätt koordinatsiooniseoses olevates identse öeldisverbiga osalausetes, kus öeldisverb on olemas ainult esimeses osalauses ning järgnevates on tüüpiliselt kustutatud. Artiklis ongi vaatluse all sellised öeldisverbi ellipsiga laused eesti keele UD puudepangas: nende lausete tüpoloogia, automaatne tuvastamine ning automaatne (ümber)märgendamine eesti keele UD puudepanga versiooni 2.2 jaoks.

1 Introduction

Universal Dependencies[1] (henceworth: UD) (McDonald et al., 2013) is an initiative aimed at developing cross-linguistically consistent treebank annotation for many languages. UD is an ongoing project, which means that the annotation guidelines and treebank annotations are subject to constant changes. Its Version 2.2 includes more than 100 treebanks in over than 70 languages.

Syntactic annotation in the UD framework represents dependency relations between tokens; the dependency arcs are labelled, i.e. they are typed dependencies. Dependency description of a clause is verb-centered: the root of a dependency tree is a finite verb form or, in copular sentences, a predicative word-form. So predicate verb ellipsis poses more problems for building a dependency tree structure than ellipsis of some argument or adjunct.

In Estonian, the most common cases of predicate verb ellipsis are gapping and stripping constructions that appear in coordinated clauses. Gapping means that the identical finite verb (together with the possible auxiliaries) is omitted in the second coordinate clause. Extended gapping construction means that some other constituent also has been elided together with the predicate verb. By stripping everything is elided from the coordinate clause except one constituent and often an additive or adversative particle is added to the single remaining constituent.

The UD version 2.2 Annotation Guidelines[2] suggest that the elliptical constructions should be annotated as follows:

1. If the elided element has no overt dependents, there is no special annotation, i.e. the elided element remains unnoticed.

2. If the elided element has overt dependents, one of these should be promoted to take the role of the head.

3. If the elided element is a predicate and the promoted element is one of its arguments or adjuncts, a special relation – orphan – should be used when attaching other non-functional dependents to the promoted head.

As Schuster et al. (2017) point out, these guidelines ”put stripping in a gray zone” as the additive/negative particle can be annotated using the relation ”orphan” or simply as an adverbial modifier.

Ellipsis has been a relatively popular research topic in UD framework. Droganova and Zeman (2017) provide an overview of annotating gapping and stripping constructions (usage of label ”orphan”) in UD 2.0 treebanks. Schuster et al. (2017) analyse gapping constructions in several languages and argue in favor of the annotation scheme of these constructions proposed in UD v2.

[1] www.universaldependencies.org

[2] http://universaldependencies.org/u/overview/specific-syntax.html#ellipsis

As the main aim of UD initiative is to facilitate multi-lingual parsing, there have already been a couple of papers that report on experiments on improving the parsing of elliptical constructions, e.g. Droganova et al. (2018) or Schuster et al. (2018).

The rest of the paper is organized as follows. Section 2 gives an overview of predicate ellipsis types in Estonian and Section 3 briefly describes their annotation in Estonian treebanks prior to Estonian UD v2.2. Method for detecting and re-annotating elliptical constructions in Estonian UD v2.2 is introduced in Section 4 as well as the main findings, i.e. types of predicate verb ellipsis in the treebank and their annotation.

2 Gapping, stripping and similar constructions in Estonian

Estonian word-order is relatively free, meaning that it is mostly determined by information structure and the main principle determining the word order is V2 (verb-second). However, there are also several clause types where the finite verb form is placed in the very beginning or in the very end (Lindström, 2017).

Gapping is norm in coordinated V2 clauses where the predicate verb has at least two dependents – the identical predicate verb is omitted in all other clauses except the first one. (Erelt, 2017, pp 598–599) So there are typically at least two orphans in an Estonian gapping clause, as in (1).

(1)	*Mari*	*sööb*	*jäätist*	*ja*	*Jüri*	*0*	*kommi.*
	Mari	eats	ice-cream-Ptv	and	Jüri	0	candy-Ptv

‘Mari is eating an ice-cream and Jüri a candy.’

An extended gapping construction is also quite common. If the clause starts with an adverbial, subject is placed next to the verb and the subject and verb are identical in coordinated clauses, the subject can be omitted together with the verb (2). (Erelt, 2017, p 599)

(2)	*Suvel*	*sööb*	*Mari*	*jäätist*	*ja*	*talvel*	*0*	*kommi.*
	Summer-Ade	eats	Mari	ice-cream-Ptv	and	winter-Ade	0	candy-Ptv

‘Mari eats ice-cream during the summer and candy during the winter.’

Non-contiguous gaps are also present in Estonian: in sentence (3) the identical finite verb form *andis* ‘gave’ and adverbial modifier *kingituseks* ‘as a gift’ are omitted in the second coordinated clause.

(3)	*Ta*	*andis*	*mulle*	*kingituseks*	*raamatu*	*ja*	*mina*	*0*	*talle*	*0*	*roosi.*
	S/he	gave	I-All	present-Trans	book-Gen	and	I	0	s/he-All	0	rose-Ptv

‘S/he gave me a book as a gift and I him/her a rose.’

The stripping construction has two subtypes: coordinating (4) and adversative (5). In both cases the elliptical clause contains only one argument plus a particle. Common additive particles in coordinating stripping constructions are *ka* ‘also’ and *samuti* ‘also’; in negative clauses *ka mitte* ‘also not’ and *samuti mitte* ‘also not’. In adversative constructions particle *mitte* ‘not’ is used.(Erelt, 2017, pp 599–601)

(4) *Jüri sööb jäätist ja Mari 0 ka.*
Jüri eats ice-cream-Ptv and Mari 0 too

'Jüri is eating an ice-cream and so does Mari.'

(5) *Jüri sööb jäätist, aga Mari 0 mitte.*
Jüri eats ice-cream-Ptv but Mari 0 not

'Jüri is eating an ice-cream but Mari is not.'

Another construction that has been annotated as an example of ellipsis in Estonian UD v 2.2 is the so-called *seda*-construction (*seda* is singular partitive case form of pronoun *see* 'it, this'), exemplified in (6).

There are two alternative ways to analyse this construction. One possibility is to consider it a clause that has undergone two successive alternations: anaphoric substitution and ellipsis. The other possibility is to interpret it as a result of one-step anaphoric substitution. In both cases, the "full" version of the sentence would look like (7). In the first alternative case, there are two alternations taking place: as the first step the whole first clause *raba pakub kordumatuid elamusi* 'marsh offers unique experiences' is substituted with anaphoric expression *teeb seda* 'does it'. As the second step, the finite verb form *teeb* 'does' is deleted. So the output of the first step – the anaphoric substitution would look like (8), which is a grammatical, well-formed Estonian sentence, and the final, elliptical version like (6).

(6) *Raba pakub kordumatuid elamusi ja seda eriti*
Mash offers unique-Pl-Ptv experience-Pl-Ptv and it-Ptv especially
talvel.
winter-Ade

'Marsh offers unique experiences, especially during the winter.'

(7) *Raba pakub kordumatuid elamusi ja raba pakub*
Mash offers unique-Pl-Ptv experience-Pl-Ptv and marsh offers
kordumatuid elamusi eriti talvel.
unique-Pl-Ptv experiences-PlPtv especially winter-Pl

(8) *Raba pakub kordumatuid elamusi ja teeb seda eriti*
Mash offers unique-Pl-Ptv experience-Pl-Ptv and does it-Ptv especially
talvel.
winter-Ade

Another way to explain the *seda*-construction is to say that the whole first clause is simply substituted with pronominal form *seda*, thus producing a verbless clause in one step.

This construction has passed unnoticed by Estonian grammar books so far, so we have no linguistic analyses to base our annotation principles on. We have decided to adopt two-step explanation (anaphora followed by ellipsis) and we treat it as an example of predicate verb ellipsis. However, this construction differs from gapping and stripping as the deleted verb is not identical with the predicate verb in the previous coordinate clause.

3 Previous treatment of ellipsis in Estonian treebanks

The Estonian UD treebank has been created by semi-automatically converting the Estonian Dependency Treebank (EDT) (Muischnek et al., 2014) into UD format. The EDT annotation scheme was based on Dependency Constraint Grammar (Karlsson et al., 1995; Bick and Didriksen, 2015).

Annotations of Estonian UD treebank are produced by several semi-automatic annotation conversions (from EDT to UD v1 and then from UD 1 to UD v2) and as such contain errors and inconsistencies.

In the original EDT the elliptical constructions were annotated so that one of the remaining arguments in the elliptical clause was promoted as the root of the clause and dependents of the deleted verb-form were annotated as dependents of the promoted argument, keeping their original syntactic labels. So the attachment of dependents is in principle same as in UD v2, but no special label (`orphan` in UD v2) was used to annotate the ”orphaned” dependents.

In contrast, UD v1 annotation scheme used a special relation `remnant` to attach dependents of the elided verb to their correlates in the coordinated clause where the verb is present. However, this annotation principle was not followed in Estonian UD v1, mainly due to lack of (human) resources for re-annotation. Figure 1 depicts annotation of an elliptical sentence in Estonian UD v1 treebank. The example sentence consists of three coordinated clauses, all sharing identical predicate verb *õpib* ‘studies’ that is omitted in the second and third clause. In the elliptical clauses the ”orphaned” subjects *Maarit* and *Ilmar* are annotated as the roots of the respective clauses and the ”orphaned” objects *kirjandust* ‘literature’ and *ajalugu* ‘history’ are, somewhat illogically, attached to the promoted subjects.

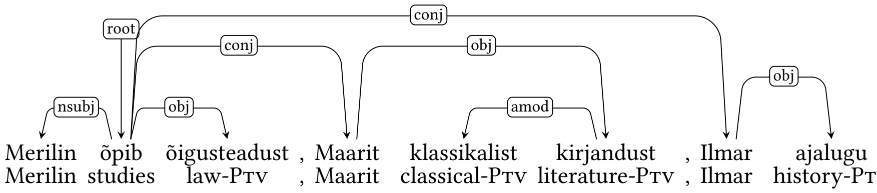
root
conj
conj
obj
obj
nsubj
obj
amod
Merilin õpib õigusteadust , Maarit klassikalist kirjandust , Ilmar ajalugu
Merilin studies law-Ptv , Maarit classical-Ptv literature-Ptv , Ilmar history-Ptv

Figure 1: ‘Merilin studies law, Maarit classical literature, Ilmar history’ (Annotation of an elliptical sentence in the Estonian UD v1 treebank)

4 Detecting and re-annotating clauses with predicate verb ellipsis in Estonian UD v2.2 treebank

As elliptical constructions were not explicitly annotated in the previous versions of the Estonian UD treebank, a special effort was needed to find and re-annotate them. A rule-based program (Python3) was created to find and re-annotate clauses with predicate verb ellipsis. The main principle is quite simple: the program looks for a verb that has a conjunct which is not a verb. This conjunct has to have at least one dependent that is not a punctuation mark or a coordinating conjunction. In order to exclude copular clauses, the conjunct also should not have a dependent labelled as copula. It means that the created piece of software works only with the locally (mis)customized version of UD v1 annotation.

As already mentioned, according to the UD v2 annotation scheme, in an elliptical clause one of the ”orphaned” dependents of the deleted verb is promoted as the head of the clause and the other dependents of the deleted verb are attached to it using the label `orphan`. However, in the enhanced version[3] of UD syntactic annotation the label `orphan` should be replaced with the label that the token would have as a dependent of the elided (and restored as a null node) verb, which is the same label that the token had in v1 of the Estonian UD treebank. In order to be able to restore the correct label in the enhanced dependencies version, we have introduced special subtypes of the label orphan, e.g `orphan:obj`, `orphan:advmod` etc.

The program achieved 95% precision and 73% recall on a test corpus consisting of 1000 sentences. It means that 95% of the detected sentences were really elliptical sentences and that 73% of the targeted sentences were actually found by the program. For the re-annotation of orphans these figures were 81.8% and 94.4%, respectively. The relatively low recall for elliptical sentence detection is mainly due to inconsistent and erroneous annotations in the treebank.

There were 359 sentences containing predicate verb ellipsis in Estonian UD treebank. Given that the treebank has a little more than 30,000 sentences, only ca 1.2% of trees contain gapping or stripping or other similar constructions.

Table 1 gives an overview of predicate verb ellipsis types in v2.2 of Estonian UD treebank as detected by the software. In the following subsections we analyse them one by one.

Type of ellipsis	Number of sentences	% of all sentences with predicate ellipsis
Simple gapping	151	42.1
Extended gapping	23	6.4
Non-contiguous gaps	19	5.3
Stripping	23	6.4
seda-construction	14	3.8
Elided copular verb	100	28
Errors	29	8
	359	100

Table 1: Frequency of predicate verb ellipsis types in Estonian UD treebank v 2.2

4.1 Simple gapping constructions

Simple gapping construction is the most frequent type of predicate ellipsis in the Estonian UD treebank, making up 42.1% of all elliptical clauses. Figures 2 and 3 depict a typical case of gapping: the identical verb-form *juhivad* ‘(they) lead’ has been deleted in the second coordinated clause, leaving an ”orphaned” object *analüüsi* ‘analysis in partitive case form’. The annotation on Figure 2 is that of v1 of Estonian UD treebank. On Figure 3 we can see the annotation that has been automatically converted into UD v2.2 format.

[3] `http://universaldependencies.org/u/overview/enhanced-syntax.html`

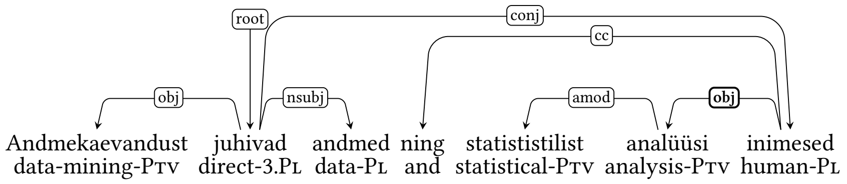
root
conj
cc
obj
nsubj
amod
obj
Andmekaevandust juhivad andmed ning statististilist analüüsi inimesed
data-mining-Ptv direct-3.Pl data-Pl and statistical-Ptv analysis-Ptv human-Pl

Figure 2: 'Data mining is directed by data and statistical analysis by humans' (Gapping construction in Estonian UD v1)

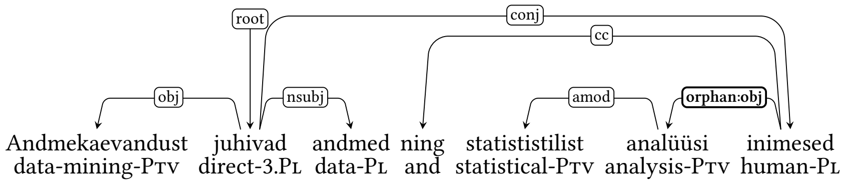
root
conj
cc
obj
nsubj
amod
orphan:obj
Andmekaevandust juhivad andmed ning statististilist analüüsi inimesed
data-mining-Ptv direct-3.Pl data-Pl and statistical-Ptv analysis-Ptv human-Pl

Figure 3: 'Data mining is directed by data and statistical analysis by humans' (Gapping construction in Estonian UD v2.2)

4.2 Extended gapping constructions

Extended gapping constructions make up 6.4% of all elliptical clauses. In this case, in addition to the predicate verb, also some of its dependents (subject, object, oblique dependents etc) are deleted. Figures 4 and 5 depict a typical case of extended gapping: both conjuncted clauses have the same finite verb *on hõivatud* 'are occupied' plus the same oblique dependent *kõrvaltöödega* 'with additional jobs', both are elided in the second clause. The annotation on Figure 4 is that of v1 of Estonian UD treebank, on Figure 5 that of the v2.2.

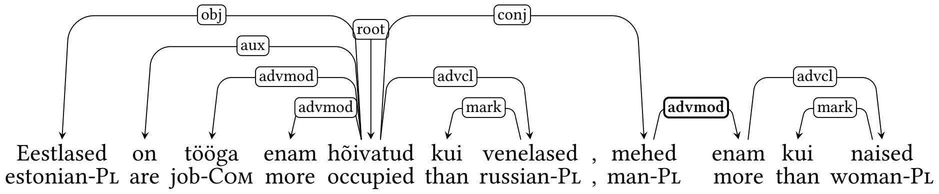
obj
root
conj
aux
advmod
advcl
advmod
mark
advmod
advcl
mark
Eestlased on tööga enam hõivatud kui venelased , mehed enam kui naised
estonian-Pl are job-Com more occupied than russian-Pl , man-Pl more than woman-Pl

Figure 4: 'Estonians are more occupied with job that Russians, men more than women.' (Extended gapping construction in Estonian UD v1)

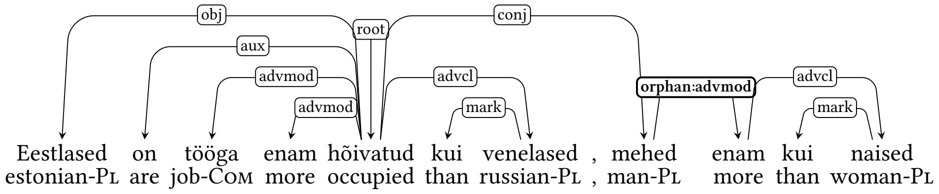
obj
root
conj
aux
advmod
advcl
advmod
mark
orphan:advmod
advcl
mark
Eestlased on tööga enam hõivatud kui venelased , mehed enam kui naised
estonian-Pl are job-Com more occupied than russian-Pl , man-Pl more than woman-Pl

Figure 5: 'Estonians are more occupied with job that Russians, men more than women.' (Extended gapping construction in Estonian UD v2.2)

4.3 Non-contiguous gaps

Non-contiguous gaps make up 5.3% of all elliptical clauses. Among them, the most frequent pattern is that the identical verb and head of a numerical or adjectival modifier are deleted from the coordinated clause.

Figures 6 and 7 depict a typical example of non-contiguous gapping: the finite verb form *maksti* 'was paid' and the adverbial modifier *rubla* 'rouble' are deleted from the coordinated clause. Deletion of finite verb leaves behind two orphaned modifiers: adverbial modifier *tollal* 'then' and oblique modifier *410* that has been promoted to the position of the deleted oblique modifier *rubla* 'rouble'.

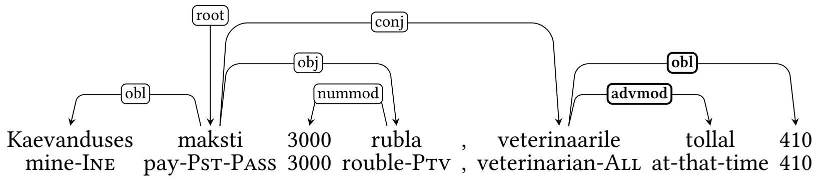
root
conj
obj
obl
nummod
obl
advmod
Kaevanduses maksti 3000 rubla , veterinaarile tollal 410
mine-Ine pay-Pst-Pass 3000 rouble-Ptv , veterinarian-All at-that-time 410

Figure 6: 'Miner was paid 3000 roubles but veterinarian 410.' (Non-contiguous gapping construction in Estonian UD v1)

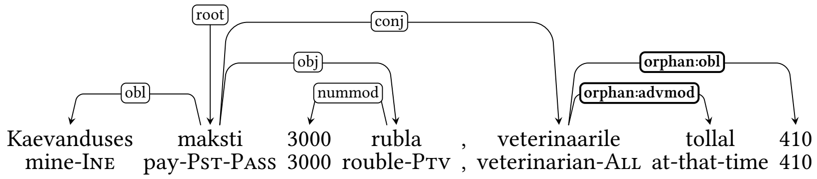
root
conj
obj
orphan:obl
obl
nummod
orphan:advmod
Kaevanduses maksti 3000 rubla , veterinaarile tollal 410
mine-Ine pay-Pst-Pass 3000 rouble-Ptv , veterinarian-All at-that-time 410

Figure 7: 'Miner was paid 3000 roubles but veterinarian 410.' (Non-contiguous gapping construction in Estonian UD v2.2)

4.4 Stripping constructions

Stripping constructions make up 6.4% of all elliptical clauses in Estonian UD treebank. As already mentioned in Section 2, by stripping everything is deleted from the coordinated clause except one argument plus an additive or adversative particle.

In the example sentences on Figure 8 and Figure 9 everything except the subject *mõned* 'some' is deleted from the coordinated clause. The adversative particle *mitte* 'not' reverses the meaning of the stripped coordinated clause. The remaining subject is annotated as the head of the clause and the adversative particle is attached to it as an adverbial modifier.

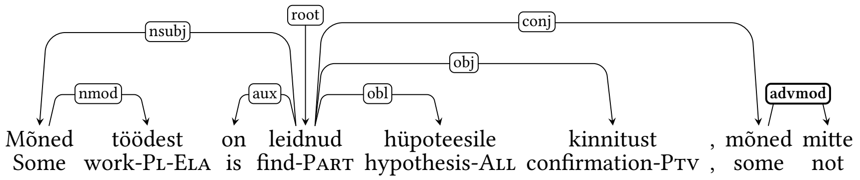
root
nsubj
conj
obj
nmod
aux
obl
advmod
Mõned töödest on leidnud hüpoteesile kinnitust , mõned mitte
Some work-Pl-Ela is find-Part hypothesis-All confirmation-Ptv , some not

Figure 8: 'Some works have confirmed the hypothesis, some not.' (Stripping construction in Estonian UD v1)

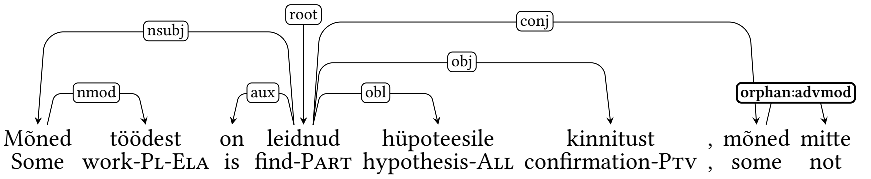
root
nsubj
conj
obj
nmod
aux
obl
orphan:advmod
Mõned töödest on leidnud hüpoteesile kinnitust , mõned mitte
Some work-PL-ELA is find-PART hypothesis-ALL confirmation-PTV , some not

Figure 9: ‘Some works have confirmed the hypothesis, some not.’ (Stripping construction in Estonian UD v2.2)

4.5 *seda*-constructions

3.8% of the clauses with predicate ellipsis are the so-called *seda*-constructions (cf Section 2). Figure 10 depicts a sentence with *seda*-construction annotated in UD v1 style and Figure 11 the same sentence as in Estonian UD v2.2. The second coordinated clause consists of pronominal form *seda* ‘it/that in partitive case form’ plus an oblique modifier *sündmusega* ‘with/by event’ and its determiner *mitme* ‘several in genitive case form’. *s*eda is annotated as the head of the elliptical clause and *s*ündmusega as its ”orphaned” oblique dependent.

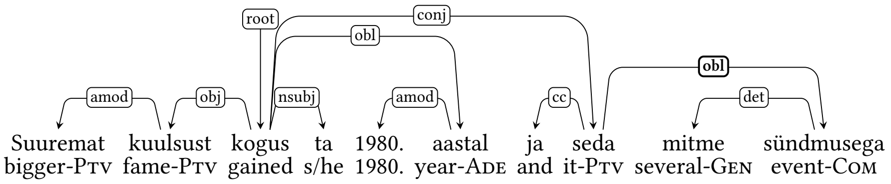
root
conj
obl
obl
amod
obj
nsubj
amod
cc
det
Suuremat kuulsust kogus ta 1980. aastal ja seda mitme sündmusega
bigger-PTV fame-PTV gained s/he 1980. year-ADE and it-PTV several-GEN event-COM

Figure 10: ‘S/he gained greater fame in the 1980s, due to several events.’ (*seda*-construction in Estonian UD v1)

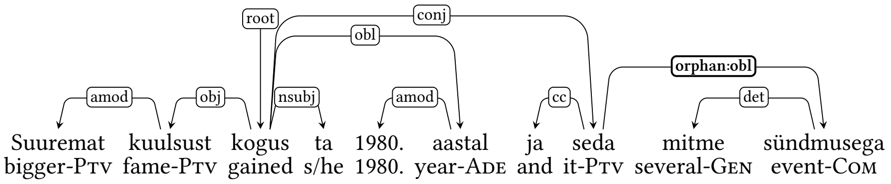
root
conj
obl
orphan:obl
amod
obj
nsubj
amod
cc
det
Suuremat kuulsust kogus ta 1980. aastal ja seda mitme sündmusega
bigger-PTV fame-PTV gained s/he 1980. year-ADE and it-PTV several-GEN event-COM

Figure 11: ‘S/he gained greater fame in the 1980s, due to several events.’ (*seda*-construction in Estonian UD v2.2)

4.6 Copular constructions

Clauses with missing copula verb *olema* ‘be’ make up 28% of all clauses with finite verb form ellipsis.

A special note on copula sentences is perhaps needed here. Our program also detects sentences with missing copular verb *olema* ‘to be’. However, according to the UD Annotation Guidelines[4], the copular clauses are regarded as instances of nonverbal predication and some argument is annotated as the root of the clause whereas the copular verb is attached to this root using the syntactic relation label cop. So, the deletion of *olema* ‘be’ does not leave behind any ”orphaned” constituents.

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

It should be pointed out that in Estonian texts there seems to be no difference between deleting *olema* or any other verb, the clause patterns are more or less the same. Figure 12 depicts a sentence with two coordinated copular clauses, where the copular verb form *on* 'is' is deleted from the second clause, but that does not result in any need for special annotation as the predicatives *kitsas* 'narrow' and *tihe* 'tight, here: heavy' serve as clause roots. The annotation is the same in both version 1 and version 2.2 of the Estonian UD treebank.

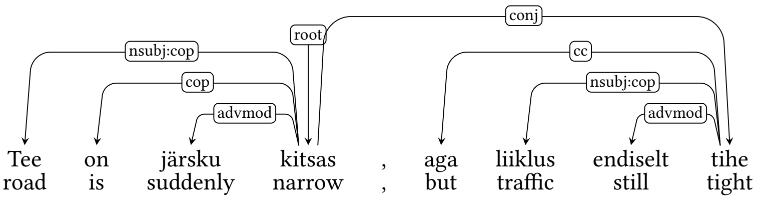

Figure 12: 'The road is suddenly narrow but traffic (is) still heavy.' (Sentence with elided copula verb in Estonian UD versions 1 and 2.2)

5 Conclusion

Predicate verb ellipsis is a difficult case for dependency syntax as the finite verb form should, as a rule, be the head of a clause. In case of its absence, dependency structure of a clause has to be constructed in some more or less artificial way.

Universal Dependencies' (UD) syntactic annotation scheme has evolved and changed over several years and so also the UD treebank annotations need to be amended and improved for the new treebank releases. Ellipsis, especially gapping and stripping, are one of those constructions that have been annotated differently in UD versions 1 and 2 and that are planned to have a special annotation (null-node insertion) in the enhanced version of UD.

This article gave an overview of predicate verb ellipsis – gapping, stripping and related constructions – in Estonian language and their frequency and annotation in the Estonian UD treebank versions 1 and 2.2.

The work described in this article has resulted in more accurate version of Estonian UD treebank. The next step would be annotating an enhanced dependencies version of the Estonian UD treebank. For elliptical constructions it means "restoring" the elided predicate verbs as null-nodes and re-attaching and re-naming its dependents.

Acknowledgments

This research has been supported by the Centre of Excellence in Estonian Studies (CEES, European Regional Development Fund) and is related to research project IUT20-56 (Estonian Research Council).

References

Eckhard Bick and Tino Didriksen. 2015. Cg-3—beyond classical constraint grammar. In *Proceedings of the 20th Nordic Conference of Computational Linguistics,* *NODALIDA 2015, May 11-13, 2015, Vilnius, Lithuania*. Linköping University Electronic Press, 109, pages 31–39.

Kira Droganova, Filip Ginter, Jenna Kanerva, and Daniel Zeman. 2018. Mind the gap: Data enrichment in dependency parsing of elliptical constructions. In *Proceedings of the Second Workshop on Universal Dependencies (UDW 2018)*. pages 47–54.

Kira Droganova and Daniel Zeman. 2017. Elliptic constructions: Spotting patterns in ud treebanks. In *Proceedings of the NoDaLiDa 2017 Workshop on Universal Dependencies (UDW 2017)*. pages 48–57.

Mati Erelt. 2017. Ellips. *Eesti keele süntaks. Tartu: Tartu Ülikool.* pages 591–601.

Fred Karlsson, Atro Voutilainen, Juha Heikkila, and Atro Anttila. 1995. *Constraint Grammar, A Language-independent System for Parsing Unrestricted Text*. Mouton de Gruyter.

Liina Lindström. 2017. Lause infostruktuur ja sõnajärg. *Eesti keele süntaks. Tartu: Tartu Ülikooli Kirjastus* pages 537–564.

Ryan McDonald, Joakim Nivre, Yvonne Quirmbach-Brundage, Yoav Goldberg, Dipanjan Das, Kuzman Ganchev, Keith Hall, Slav Petrov, Hao Zhang, Oscar Täckström, et al. 2013. Universal dependency annotation for multilingual parsing. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. volume 2, pages 92–97.

Kadri Muischnek, Kaili Müürisep, Tiina Puolakainen, Eleri Aedmaa, Riin Kirt, and Dage Särg. 2014. Estonian dependency treebank and its annotation scheme. In *Proceedings of 13th Workshop on Treebanks and Linguistic Theories (TLT13)*. pages 285–291.

Sebastian Schuster, Matthew Lamm, and Christopher D Manning. 2017. Gapping constructions in universal dependencies v2. In *Proceedings of the NoDaLiDa 2017 Workshop on Universal Dependencies (UDW 2017)*. pages 123–132.

Sebastian Schuster, Joakim Nivre, and Christopher D. Manning. 2018. Sentences with gapping: Parsing and reconstructing elided predicates. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2018, New Orleans, Louisiana, USA, June 1-6, 2018, Volume 1 (Long Papers)*. pages 1156–1168. https://aclanthology.info/papers/N18-1105/n18-1105.