

Enhanced Universal Dependencies in the Wild: Evaluating Portuguese EUD Parsing in Realistic Scenarios

Elvis A. de Souza

Núcleo Interinstitucional de
Linguística Computacional (NILC)
Universidade de São Paulo
elvis.desouza99@gmail.com

Thiago A. S. Pardo

Núcleo Interinstitucional de
Linguística Computacional (NILC)
Universidade de São Paulo
taspardo@icmc.usp.br

Abstract

Enhanced Universal Dependencies (EUD) provide a more informative syntactic representation than Basic Universal Dependencies by relaxing tree constraints to allow for graph structures and the inclusion of other relationships. While conversion rules from basic to enhanced relations have been established for Portuguese, they were previously evaluated only on journalistic text using gold-standard basic syntactic trees as input. This paper evaluates the robustness of these rules in diverse scenarios (“in the wild”), encompassing other text genres and domains, as well as realistic parsing pipelines to automatically generate the input data. Our results demonstrate that Portuguese-specific rules consistently outperform universal rules. However, the reliance on automatic basic syntax significantly impacts performance. This degradation is particularly severe when the domain of the input text differs from the training data of the basic parser. We also provide a detailed error analysis, identifying specific difficult linguistic phenomena and scenarios.

1 Introduction

Syntactic parsing is a core NLP task for understanding linguistic structure, important for areas such as social sciences and computational linguistics (Jurafsky and Martin, 2025). It structures data for processing and aids in interpreting neural model outputs (Jurafsky and Martin, 2025). As an intermediate step, parsing enhances information extraction, question answering, and summarization, among other tasks (Grishman, 1997; Di Caro and Grella, 2013; Wu et al., 2018; Song et al., 2019; Alonso et al., 2021; Freitas et al., 2023). Beyond the development of better parsers, research also focuses on richer linguistic representations like Enhanced Universal Dependencies (EUD), which bridge syntax and semantics (Schuster and Manning, 2016).

EUD is an optional layer of the Universal Dependencies (UD) project, which provides a multilin-

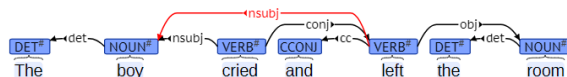


Figure 1: Syntactic annotation of the sentence “The boy cried and left the room”, in which the red arc is an enhanced relation.

gual grammar framework for over 150 languages (Nivre et al., 2016; de Marneffe et al., 2021). Despite its utility, EUD remains relatively unexplored compared to basic UD syntactic annotation (Nivre et al., 2016; Schuster and Manning, 2016; Bouma et al., 2020, 2021). By relaxing the formal constraints of traditional syntactic trees, EUD produces enriched graphs (Schuster and Manning, 2016). For instance, Figure 1 shows an EUD arc (in red) where the token “boy” depends on two clauses. This indicates that coordinated clauses share a subject—a more informative analysis that violates traditional tree constraints where a token must have exactly one governor.

In 2024, de Souza et al. (2024) conducted the first experiments on EUD annotation for Portuguese. The authors fine-tuned linguistic conversion rules from Basic Universal Dependencies (BUD), the traditional UD syntactic analysis, to EUD, the enriched analysis, on a Portuguese treebank consisting of journalistic texts included in the UD catalog (Duran et al., 2023). The experiments achieved results of up to 96.05% F₁-score when starting from gold-standard basic syntactic annotation.

The present work aims to evaluate a similar set of rules in real, more diverse, and more challenging scenarios. Because the BUD → EUD conversion rules were fine-tuned on a treebank of the journalistic domain, we also assess their performance on a legal-domain treebank (Lopes et al., 2025), on a treebank consisting of user-generated content from social media (Di-Felippo and Roman, 2025),

and on a literary treebank composed of “The Little Prince” book. Furthermore, in real-world scenarios where a system is deployed and receives new data on a daily basis, gold-standard basic syntactic annotation is not available. For this reason, we also evaluate EUD parsing performance when relying on automatically produced BUD generated by a state-of-the-art parser for Portuguese (Lopes and Pardo, 2024).

This paper is organized as follows. Section 2 brings the related work. Section 3 describes our datasets. Section 4 details our experiments, while evaluation and results are shown in Section 5. Section 6 concludes this paper.

2 Related work

There is a series of works discussing the relevance of EUD annotation and its use in treebanks (Nivre et al., 2016; Schuster and Manning, 2016; Nivre et al., 2020). The instantiation of EUD for Portuguese was introduced by Pagano et al. (2023) and also addressed in other works (de Souza et al., 2024; Duran, 2024; Duran et al., 2025).

Overall, EUD graphs may include six phenomena that are not possible in BUD:

1. Inclusion of the lemmas of prepositions, coordinating conjunctions, and subordinating conjunctions in the label of the relations they introduce (as in Figure 2);
2. Identification of the controlling subject of the null subject in open clausal complements (*xcomp* clauses) (Figure 3);
3. Propagation to the dependent of *conj* of the relation that targets the head of *conj* (Figure 4);
4. Propagation to the dependent of *conj* of certain relations that depart from the head of *conj* (Figure 1);
5. Replacement of the relative pronoun in relative clauses with its antecedent, while marking the relation between the pronoun and its antecedent with the EUD-specific label *ref* (Figure 5);
6. Insertion of an empty token to represent an elliptical predicate and establishment of relations between this empty token and the participants of the *orphan* relation (Figure 6).

The shared tasks held at IWPT in 2020 (Bouma et al., 2020) and 2021 (Bouma et al., 2021) provided a platform for comparing results among different systems for EUD parsing from raw text. The

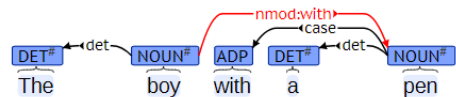


Figure 2: EUD annotation, where the relation *nmod* (nominal modifier) is extended with the word “with”.

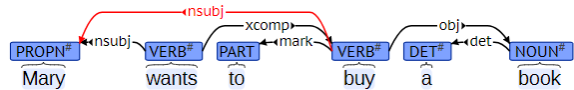


Figure 3: EUD annotation showing an *nsubj* (subject) governed by a verb dependent on *xcomp* (open clausal complement).

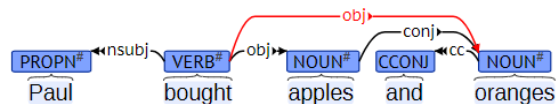


Figure 4: EUD annotation, where the *obj* (object) relation is propagated to the dependent of *conj* (conjunct).

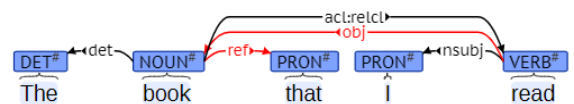


Figure 5: EUD annotation, where “book” is the object of “read” and “that” is the *ref* of “book”.

best-performing systems in 2021 employ a wide variety of architectures and approaches, including different preprocessing tools, rule-based or neural-based approaches, graph-based or tree-based parsing with post-processing steps, and so on. The shared tasks evaluated results for 17 different languages; Portuguese was not among them, as no EUD-annotated data was available at the time.

In de Souza et al. (2024), the authors adapted the rules from a rule-based annotation system, named Grew, for Portuguese. The system original submission to the IWPT 2021 shared task ranked seventh, with 81.58% F₁-score (an average of the 17 participant languages), and was the highest-ranked symbolic system.¹ The adaptation for Portuguese resulted in fifteen new rules, an increase of 12.5% in the number of mapped rules, in addition to several minor modifications to existing rules. The modified rules were tested on a held-out human-annotated sample of the treebank used for the development of the rules, Porttinari-base (Duran et al., 2023), a treebank from the journalistic domain with gold-standard basic syntactic annotation. The authors observed an increase of 11.38 percentage points in F₁-score when using the Portuguese-tuned rules to

¹Grew is 7.66 points below the top-ranked system, TGIF (Shi and Lee, 2021), which is neural-based.

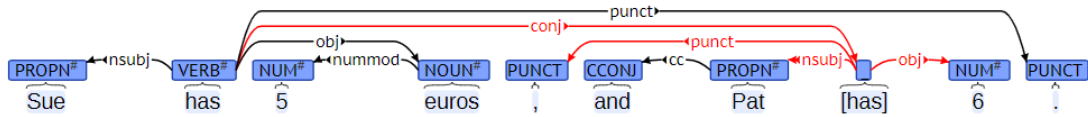


Figure 6: EUD annotation for the sentence “Sue has 5 euros, and Pat 6.”, in which an empty token *[has]* is inserted.

annotate the Portuguese dataset compared with the original rules submitted to IWPT 2021, reaching 96.05% F₁-score.

Since graph rewriting rules depend on the basic syntactic tree, defective basic syntax is expected to significantly reduce the performance of automatic EUD annotation. For example, while the Grew submission scored only 81.58% F₁-score in the official joint evaluation, which used automatically produced basic syntax, Guillaume and Perrier (2021) tested the system using gold-standard basic syntax and obtained an average F₁-score of 97%.

This work fills two gaps left open for Portuguese: we assess the impact that automatic BUD annotation may have on EUD parsing, and we evaluate the same rules when applied to other text genres/domains.

3 Data

The treebanks selected to assess the robustness of EUD parsing for Portuguese show distinct linguistic features. All treebanks contain gold-standard basic syntactic UD annotation, meaning they share the same annotation guidelines and can be directly processed by an EUD annotation system that relies on basic syntax. With the exception of Porttinari-base, which is already available with EUD annotation, the remaining treebanks do not contain gold-standard EUD annotation. It was therefore necessary to create such annotation samples for the purposes of this study.

The gold-standard EUD annotation samples were developed following the same criteria used for the Porttinari-base reference annotation (de Souza et al., 2024). A set of 100 randomly selected sentences was extracted from the test sets of the treebanks. The goal was to get a reduced sample that still reflected the distribution of linguistic phenomena present in the full treebanks. The gold annotation was created by manually correcting the output of the automatic EUD annotation system. This review process was carried out by a single annotator, expert in both basic and enhanced Universal Dependencies syntax.

The treebanks and their test data sampled for the

EUD annotation can be described as follows.

Porttinari-base (Duran et al., 2023) This treebank contains 8,418 sentences and 168,080 tokens from journalistic texts. The sample of 100 test sentences selected in de Souza et al. (2024) and also used here contains 2,012 tokens, an average of 20.12 tokens per sentence, and a Type-Token Ratio of 0.3921. The treebank contains sentences such as the following: *Kim Jong-un tomou as rédeas do país com a morte de seu pai, em dezembro de 2011.*²

PortJur (Lopes et al., 2025) This treebank contains 2,005 sentences and 101,297 tokens from the legal domain, with texts from public law produced by the judiciary (mainly summaries) and the legislature (laws). It includes widely known Brazilian laws such as the Internet Civil Rights Law and Maria da Penha Law, among many others. The sample of 100 sentences used for EUD annotation contains 4,706 tokens, with an average of 47.06 tokens per sentence. This is the highest average among the treebanks, with a number of tokens 2.33 times larger than the second largest, Porttinari-base. In addition, it has a Type-Token Ratio of 0.2654. Thus, although its sentences are long, they are considerably less lexically diverse than those in the other treebanks. The treebank contains sentences such as: *Havendo danos às pessoas idosas abrigadas ou qualquer tipo de fraude em relação ao programa, caberá o afastamento provisório dos dirigentes ou a interdição da unidade e a suspensão do programa.*³

DANTEStocks (Di-Felippo and Roman, 2025) This treebank contains 4,042 user-generated posts (from Twitter/X) and 80,997 tokens from the financial domain. The test sample, composed of 100 random posts, contains 1,958 tokens, an average of 19.58 tokens per post, and a Type-Token Ratio

²“Kim Jong-un took the reins of the country after the death of his father in December 2011.”

³“If there is harm to the elderly residents or any type of fraud related to the program, the temporary removal of the directors or the closure of the unit and the suspension of the program will be warranted.”

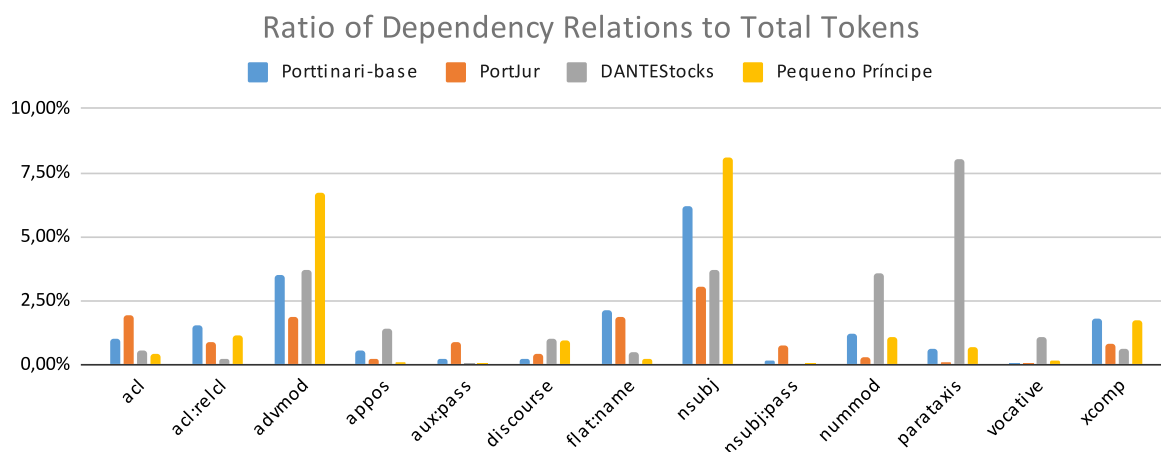


Figure 7: Characterization of the treebank samples by the number of dependency relations in the basic UD trees.

of 0.3825, both values close to those of Porttinari-base. The treebank contains sentences such as: *@danielschettin0 Tô contigo... VALE5 tem um peso MONSTRUOSO na minha carteira, por causa da venda coberta de opções. abc!*⁴

Pequeno Príncipe This treebank contains 1,527 sentences and 16,802 tokens from the book “The Little Prince”, which is in the public domain.⁵ The test sample consists of 100 sentences and 1,148 tokens, an average of 11.48 tokens per sentence and a Type-Token Ratio of 0.2936. While it is the smallest of the four datasets, its lexical diversity is higher only than that of PortJur. The treebank contains sentences such as: *Todas as estrelas estão floridas.*⁶

Another way to contrast the treebanks is by examining the number of basic syntactic relations they contain. Figure 7 shows the frequency of dependency relations in the samples of each treebank, proportionally to the total number of syntactic dependencies. Only dependency relations that reveal differences among the treebanks were selected.

Porttinari-base has higher number of subject relations (*nsubj*, 6.21 percent of all relations), relative clauses (*acl:relcl*, 1.54 percent), and open clausal complements (*xcomp*, 1.79 percent). Compared with PortJur and DANTEStocks, it contains more sentences with full clausal structure and overt subjects.

⁴@danielschettin0 I’m with you... VALE5 has a HUGE weight in my portfolio because of the covered call options. Cheers!”

⁵The annotated data can be obtained from: <https://sites.google.com/icmc.usp.br/poetisa/porttinari-3-0>.

⁶“All the stars are in bloom.”

PortJur, in contrast, contains more *acl* relations (1.91 percent) than *acl:relcl* relations (0.91 percent), unlike Porttinari-base. This may indicate a greater number of reduced clauses than fully developed relative clauses. Compared with the other treebanks, PortJur has considerably fewer adverbial modifiers (*advmod*, 1.87 percent), a high number of passive constructions (*aux:pass*, 0.91 percent, and *nsubj:pass*, 0.76 percent), a mark of formal and complex writing style, and a low number of numeral modifiers (*nummod*, 0.28 percent) and *parataxis* relations (0.08 percent).

DANTEStocks, in turn, has few fully developed relative clauses (*acl:relcl*, 0.26 percent), a high number of appositions (*appos*, 1.38 percent), *discourse* relations (1.02 percent), numeral modifiers (*nummod*, 3.58 percent), *parataxis* relations (8.02 percent), and *vocative* relations (1.07 percent), indicative of informal writing. Proper names tend not to be multiword, as indicated by the low rate of *flat:name* relations (0.46 percent).

Pequeno Príncipe is similar to Porttinari-base in many syntactic aspects, such as the number of fully developed relative clauses (*acl:relcl*, 1.13 percent), numeral modifiers (*nummod*, 1.05 percent), *parataxis* (0.70 percent) and open clausal complements (*xcomp*, 1.74 percent). In relations such as adverbial modifiers and subjects, it follows the same trend of Porttinari-base, but surpasses it (*advmod*, 6.71 percent, and *nsubj*, 8.10 percent).

Table 1 describes the composition of the test samples from the treebanks, whose enhanced annotation was manually reviewed to ensure that they are gold-standard annotations. According to the EUD guidelines, BUD relations that are not af-

ected by any of the EUD enrichment rules should be copied from the BUD trees to the EUD graphs, but we show in this table only the relations that are different from those in BUD to facilitate interpretation. As a consequence, the number of EUD relations that are shown is lower than the number of tokens in the samples.

	Portt.	PortJur	Dant.St.	P.Prínc.
Sentences	100	100	100	100
Tokens	2,012	4,706	1,958	1,148
EDepRel.	554	1,813	398	184
EUD 1	347	1,130	241	116
EUD 2	33	41	5	17
EUD 3	55	232	111	10
EUD 4	26	169	22	11
EUD 5	55	81	7	24
EUD 6	10	0	4	0

Table 1: Description of the gold-standard EUD test samples from the treebanks. EDepRel is the number of enhanced syntactic dependencies that are different from those in BUD trees; EUD 1–6 indicate the number of relations of each EUD type.

Figure 8 complements Table 1 by showing the proportion of each EUD type within the total number of enhanced relations in each treebank. EUD type 1, assigned to functional words, is the most frequent across all treebanks, consistently accounting for over 60% of the relations.

The distribution of the subsequent types reveals distinct syntactic profiles: in PortJur and DANTEStocks, EUD type 3 (propagation of governors in coordination) is the second most frequent, especially in DANTEStocks, where it reaches 27.89%. This highlights the importance of coordination phenomena in financial texts, such as the list of financial assets. In Porttinari-base, the second position is shared by EUD type 3 and EUD type 5 (relative clauses), both at 9.93%. EUD type 2 (*xcomp* relations) follows these with 5.96%. Pequeno Príncipe stands out for the high frequency of EUD type 5 (13.04%), which is its second most common type, followed by EUD type 2 (9.24%). EUD type 4, which propagates dependents of coordinated elements, is particularly relevant in PortJur (9.32%), whereas it is less frequent in the others. Finally, relations involving predicate ellipsis (EUD type 6) are negligible or entirely absent, as seen in PortJur and Pequeno Príncipe.

Figure 9 illustrates the increase in selected syntactic relations within EUD graphs compared to

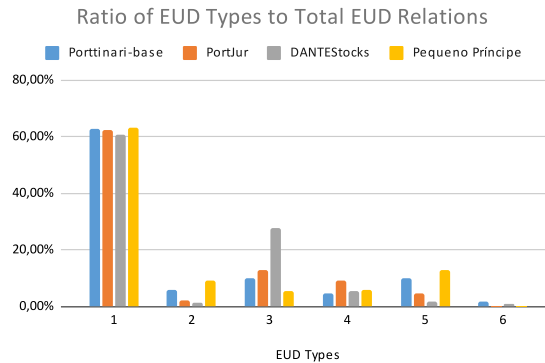


Figure 8: EUD relations by type in each test sample.

basic UD trees (using gold-standard annotations). When compared to Figure 7, these results demonstrate that treebanks differ not only in their basic relation proportions but also in how EUD rules multiply those occurrences across different datasets.

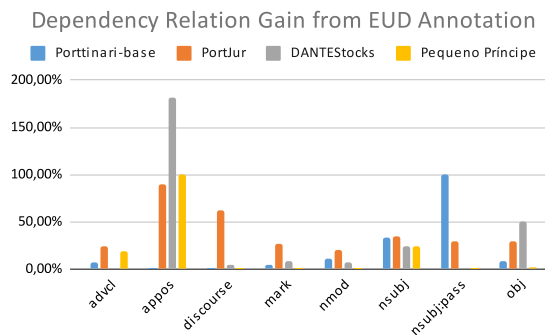


Figure 9: Increase in relations in the EUD graph of the test samples compared to the basic UD trees.

The relation *appos* is more frequent in DANTEStocks (1.38%) than in Porttinari-base (0.55%), PortJur (0.21%) or Pequeno Príncipe (0.09%) (Figure 7); however, the increase in *appos* relations in the PortJur and Pequeno Príncipe EUD graphs are unexpectedly high, having 90% and 100% more *appos* relations in the EUD graphs than in their BUD trees, respectively, second only to the increase observed in DANTEStocks (181.48%). Porttinari-base, in contrast, shows no increase in *appos* relations after EUD annotation (Figure 9).

Relations such as *discourse*, assigned mostly to discourse elements such as bullet points in legal texts and markers of orality in social media text, are more frequent in DANTEStocks (1.02% of basic relations) and in Pequeno Príncipe (0.96%) than in PortJur (0.45%) or Porttinari-base (0.20%). However, the increase in *discourse* relations in the EUD graphs of DANTEStocks and Pequeno Príncipe are

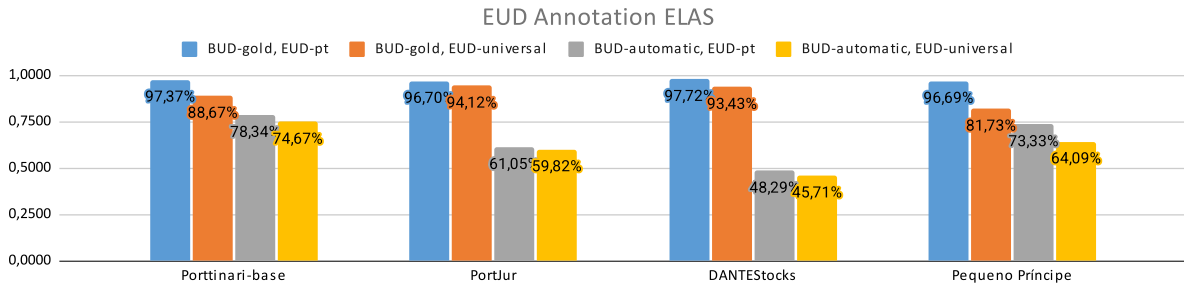


Figure 10: EUD annotation ELAS for the different experiment configurations.

5% and 0%, respectively, while PortJur shows an increase of 61.90%. Porttinari-base, which had the fewest discourse relations, shows no increase in the EUD graph.

Relations such as *nsubj:pass*, for passive subjects, are more frequent in PortJur (0.76%) than in Porttinari-base (0.15%), DANTEStocks (0.05%) or Pequeno Príncipe (0%). However, the increase in passive subject relations is larger in Porttinari-base (100%) than in PortJur (30.08%). DANTEStocks shows no increase.

4 Experiments

The automatic EUD annotation system used in this work is Grew (Bonfante et al., 2018; Guillaume and Perrier, 2021), a graph-rewriting system (see Section 2). The system and its rules developed for the shared task in 2021 are publicly available.

The adapted rules for Portuguese are available in the repository cited by de Souza et al. (2024). By checking the updates in the repository, we realize that several improvements have been made to the rules since the original publication. For this reason, we also replicate the authors’ work for the Porttinari-base treebank, but using the newer rules.

We compare the performance of different rule sets applied to different input data.

1. Regarding the rule sets, we evaluate both the rules developed for IWPT 2021 (Guillaume and Perrier, 2021) and the mentioned updated rules for EUD annotation in Portuguese. The first rule set is referred to as *EUD-universal*, because it was designed to be applicable to any of the languages participating in the shared task (Portuguese was not a participant because it lacked annotated resources). The second rule set is referred to as *EUD-pt*, since these rules were adapted for Portuguese.
2. Regarding the input data, variations occur

at two levels: (i) as described in Section 3, we evaluate Portuguese treebanks from different text genres/domains, with diverse linguistic properties; (ii) at the level of basic syntactic annotation, each treebank is tested in two versions—one with gold-standard basic UD annotation, called *BUD-gold*, and another with basic UD annotation produced by a state-of-the-art Portuguese parser, Portparser v2 (Lopes and Pardo, 2024), called *BUD-automatic*.

5 Evaluation and results

The metric used to evaluate the system is ELAS, labeled attachment score on enhanced dependencies. ELAS is an F_1 -score computed over the set of predicted and reference EUD. In addition to the overall ELAS, we also compute ELAS by EUD type, following the categories in Section 2.

Figure 10 presents the performance of the EUD annotation system in different scenarios: (a) when applied to Porttinari-base, PortJur, DANTEStocks and Pequeno Príncipe; (b) when applied to gold-standard basic syntax (*BUD-gold*, blue and orange bars); (c) when applied to automatic basic syntax (*BUD-automatic*, gray and yellow bars); (d) when using the Portuguese adapted rules (*EUD-pt*, blue and gray bars); and (e) when using the universal IWPT 2021 rules (*EUD-universal*, orange and yellow bars).

Although the treebanks differ considerably in terms of basic syntax (Figure 7) and in how EUD rules apply to them (Figure 9), a similar performance pattern holds across all four. The Portuguese EUD rules consistently outperform the universal rules. The largest gap is found in Pequeno Príncipe, where performance increases by 14.96 points.

The performance gap is larger when comparing application over gold-standard versus automatic basic syntax. The difference is sharpest in DAN-

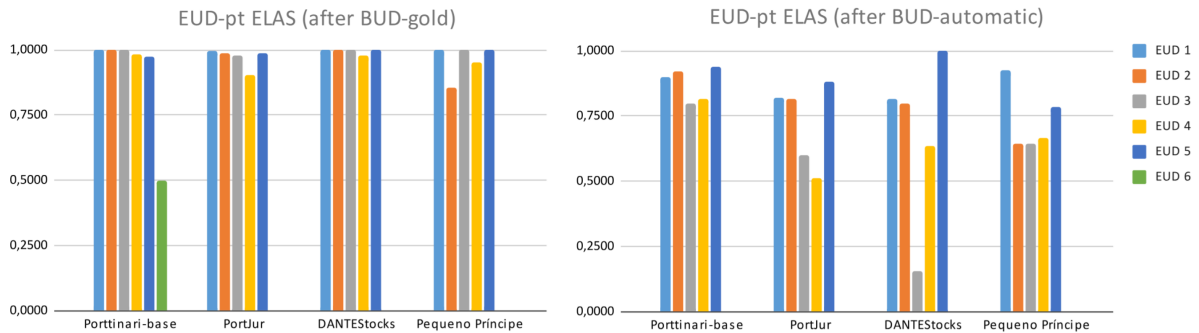


Figure 11: ELAS for each EUD type using the EUD-pt rules, with BUD-gold and BUD-automatic.

TEStocks, where the difference between BUD-gold/EUD-pt and BUD-automatic/EUD-pt reaches 49.43 points. This is mainly due to the fact that Portparser, the basic dependency parser used for automatic annotation, was trained on Porttinari-base (Lopes and Pardo, 2024). Supporting this hypothesis, the difference is substantially smaller for Porttinari-base, at only 19.03 points.

Figure 11 provides a detailed view of the performance of each EUD type when applying the Portuguese rules to the treebanks with gold-standard basic syntax (BUD-gold/EUD-pt, bars on the left) and with automatic basic syntax (BUD-automatic/EUD-pt, bars on the right). Although overall results followed the same trend in Figure 10, Figure 11 shows that the EUD types with highest and lowest performance differ across treebanks.

EUD type 6, assigned to predicate ellipsis, is nearly absent from the test samples. All annotations involving predicate ellipsis were incorrect when the rules were applied over the automatic basic syntax. Because this phenomenon is extremely rare in the four samples, this work cannot contribute new conclusions about its automatic annotation performance. The conclusions from de Souza et al. (2024) remain valid: predicate ellipsis is a difficult phenomenon to annotate automatically using a rule-based approach.

For the remaining EUD types, when the rules were applied to gold-standard basic syntax, results were above 97% for all types except EUD type 4, which reached 90.4% in PortJur and 95.2% in Pequeno Príncipe, and type 2, which reached 85.7% in Pequeno Príncipe. Most errors of EUD type 4 resemble those in Figure 12, where coordination occurs within lists marked by bullet points. In such cases, in the gold-standard BUD, the bullet points are annotated as dependents of a *discourse* relation governed by an item to their right (“prática”, “fi-

nalidade”, “participação”). Since these governors are coordinated, the EUD type 4 rule propagates the dependent element (the bullet point “(i)”) to all coordinated tokens, which is incorrect.

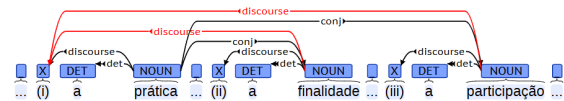


Figure 12: Excerpt from PortJur (BUD-gold) showing an EUD type 4 annotation error.

With automatic basic syntax as input, results were more diverse. For EUD type 1, performance was higher in Pequeno Príncipe (92.8%) and Porttinari-base (90.0%) and lower in PortJur and DANTEStocks (82.1% and 81.5%, respectively). An example from DANTEStocks is shown in Figure 13. The automatic annotation assigned a compound proper noun (*flat:name*) structure to “Lancamento De as Notas Seniores”. Because of this, the expected *nmod:de* relation between “Notas” and “Lancamento” did not appear in the EUD graph.

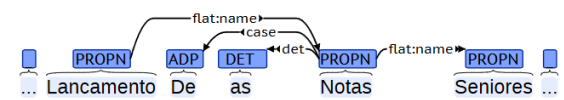


Figure 13: Excerpt from DANTEStocks (BUD-automatic) showing an EUD type 1 annotation error.

For EUD type 2, results are 92.3% for Porttinari-base, 81.4% for PortJur, 80.0% for DANTEStocks and 64.5% for Pequeno Príncipe. Figure 14 shows an example from DANTEStocks. The token “PETR4”, at the beginning of the sentence, was incorrectly annotated in the automatic basic syntax as the subject of the main verb “deverá”. This verb is the governor of an *xcomp* relation with “decidir”. After applying the EUD rules, PETR4 became the

subject of “decidir” as well, which is incorrect. In the gold-standard annotation PETR4 is annotated only as a *parataxis* of “deverá”, not as a subject.

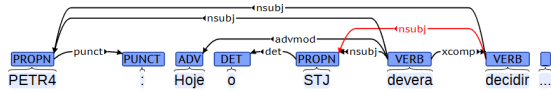


Figure 14: Excerpt from DANTEStocks (BUD-automatic) showing an EUD type 2 annotation error.

For EUD type 3, performance drops sharply in DANTEStocks, with only 15.4% accuracy, compared to 80.0%, 60.0% and 64.3% in Porttinari-base, PortJur and Pequeno Príncipe, respectively. In Figure 15, the automatic syntactic annotation contains two errors: it annotates “SEER3” and “SMTO3” as *nmod* dependents of “SANB11” instead of coordinated tokens, and it annotates “SANB11” as the dependent of a *parataxis* instead of *appos*. In the gold-standard EUD annotation, the *appos* relation between “Ativo” and “SANB11” should be propagated to the other assets, “SEER3” and “SMTO3”, which is impossible due to the two errors.

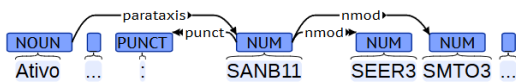


Figure 15: Excerpt from DANTEStocks (BUD-automatic) showing an EUD type 3 annotation error.

For EUD type 4, PortJur had the lowest performance (51.0%), compared to 81.3%, 63.4% and 66.7% for Porttinari-base, DANTEStocks and Pequeno Príncipe, respectively. Most cases resemble those shown in Figure 12, since the same type of error occurs in both BUD-gold and automatic annotations.

For EUD type 5, DANTEStocks achieved perfect performance (100%), while Porttinari-base scored 93.7%, followed by PortJur with 88.3% and Pequeno Príncipe with 78.2%. Figure 16 shows an example from PortJur. The relative clause headed by “identificada” was incorrectly annotated in the BUD tree as a dependent of “agravo”, a token located approximately 60 tokens to the left (in a sentence of 244 tokens) and which does not even agree morphologically with the verb. It should have been annotated as a dependent of “propaganda”. Consequently, the relative pronoun “que” was incorrectly

annotated in the EUD graph as referring to “agravo” instead of “propaganda”.

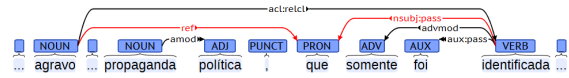


Figure 16: Excerpt from PortJur (BUD-automatic) showing an EUD type 5 annotation error.

6 Final remarks

This work evaluated the robustness of a rule-based system for Enhanced Universal Dependencies (EUD) annotation in Portuguese, expanding previous research to cover diverse textual genres and realistic parsing scenarios. By testing on legal (PortJur), social media (DANTEStocks) and literary (Pequeno Príncipe) texts, in addition to the journalistic domain (Porttinari-base), we demonstrated that the EUD conversion rules adapted specifically for Portuguese consistently outperform the universal rule set originally developed for the IWPT 2021 shared task. This confirms the importance of language-specific fine-tuning for rule-based syntactic enrichment.

The transition from gold-standard to automatically generated basic syntax revealed significant challenges for “in the wild” applications. Our experiments showed that the performance of the rewriting rules is heavily dependent on the quality of the input Basic Universal Dependencies (BUD) tree. While results remain high when the basic parser performs well, performance degrades sharply in out-of-domain scenarios. This was particularly evident in the DANTEStocks corpus, where the basic parser, trained on standard journalistic text, struggled with user-generated content, causing a cascade of errors in the enhanced representation.

These findings suggest that, while the current rule-based approach is highly effective for converting high-quality syntactic trees, its utility in fully automated pipelines is constrained by the robustness of the underlying basic parser. Future work should focus on two complementary directions: improving the domain adaptation of basic Portuguese parsers to reduce upstream errors, and expanding the coverage of gold-standard EUD treebanks for Portuguese. The reference samples produced in this study serve as a step toward enabling the training of robust, fully neural EUD parsers that may

be less sensitive to the strict constraints of basic syntactic trees.

The interested reader may find more information about this work at the web portal of the POeTiSA project⁷.

Acknowledgments

This work was carried out at the Center for Artificial Intelligence of the University of São Paulo (C4AI - <http://c4ai.inova.usp.br/>), with support by the São Paulo Research Foundation (FAPESP grant #2019/07665-4) and by the IBM Corporation. The project was also supported by the Ministry of Science, Technology and Innovation, with resources of Law N. 8,248, of October 23, 1991, within the scope of PPI-SOFTEX, coordinated by SofTex and published as Residence in TIC 13, DOU 01245.010222/2022-44, and by *Instituto Nacional de Ciência e Tecnologia em Inteligência Artificial Responsável para Linguística Computacional, Tratamento e Disseminação de Informação* (INCT-TILD-IAR) (grant #408490/2024-1).

References

- Miguel A. Alonso, Carlos Gómez-Rodríguez, and Jesús Vilares. 2021. [On the Use of Parsing for Named Entity Recognition](#). *Applied Sciences*, 11(3).
- Guillaume Bonfante, Bruno Guillaume, and Guy Perrier. 2018. *Application of Graph Rewriting to Natural Language Processing*. John Wiley & Sons.
- Gosse Bouma, Djamé Seddah, and Daniel Zeman. 2020. [Overview of the IWPT 2020 Shared Task on Parsing into Enhanced Universal Dependencies](#). In *Proceedings of the 16th International Conference on Parsing Technologies and the IWPT 2020 Shared Task on Parsing into Enhanced Universal Dependencies*, pages 151–161, Online. Association for Computational Linguistics.
- Gosse Bouma, Djamé Seddah, and Daniel Zeman. 2021. [From Raw Text to Enhanced Universal Dependencies: The Parsing Shared Task at IWPT 2021](#). In *Proceedings of the 17th International Conference on Parsing Technologies and the IWPT 2021 Shared Task on Parsing into Enhanced Universal Dependencies (IWPT 2021)*, pages 146–157, Online. Association for Computational Linguistics.
- Marie-Catherine de Marneffe, Christopher D Manning, Joakim Nivre, and Daniel Zeman. 2021. Universal dependencies. *Computational linguistics*, 47(2):255–308.
- Elvis de Souza, Magali Duran, Maria das Graças Nunes, Gustavo Sampaio, Giovanna Belasco, and Thiago Pardo. 2024. [Automatic Annotation of Enhanced Universal Dependencies for Brazilian Portuguese](#). In *Anais do XV Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana*, pages 217–226, Porto Alegre, RS, Brasil. SBC.
- Luigi Di Caro and Matteo Grella. 2013. Sentiment analysis via dependency parsing. *Computer Standards & Interfaces*, 35(5):442–453.
- Ariani Di-Felippo and Norton Trevisan Roman. 2025. DANTEstocks: a multi-layered annotated corpus of stock market tweets for Brazilian Portuguese. *Revista Brasileira de Linguística Aplicada*, 25(1):e49802.
- Magali Duran, Lucelene Lopes, Maria das Graças Nunes, and Thiago Pardo. 2023. [The Dawn of the Portinari Multigenre Treebank: Introducing its Journalistic Portion](#). In *Anais do XIV Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana*, pages 115–124, Porto Alegre, RS, Brasil. SBC.
- Magali S. Duran, Elvis A. de Souza, Maria das Graças Volpe Nunes, Adriana Silvina Pagano, and Thiago A. S. Pardo. 2025. [Extending the Enhanced Universal Dependencies – addressing subjects in pro-drop languages](#). In *Proceedings of the Eighth Workshop on Universal Dependencies (UDW, SyntaxFest 2025)*, pages 143–152, Ljubljana, Slovenia. Association for Computational Linguistics.
- Magali Sanches Duran. 2024. [Anotação de enhanced dependencies](#). São Carlos: Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo.
- Cláudia Freitas, Elvis de Souza, Maria Clara Castro, Tatiana Cavalcanti, Patricia Ferreira da Silva, and Fábio Corrêa Cordeiro. 2023. Recursos linguísticos para o PLN específico de domínio: o Petrolês. *Linguamática*, 15(2):51–68.
- Ralph Grishman. 1997. Information Extraction: Techniques and Challenges. In *International Summer School on Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology*, SCIE '97, page 10–27, Berlin, Heidelberg. Springer-Verlag.
- Bruno Guillaume and Guy Perrier. 2021. Graph Rewriting for Enhanced Universal Dependencies. In *IWPT 2021-17th International Conference on Parsing Technologies*.
- Daniel Jurafsky and James H. Martin. 2025. [Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, with Language Models](#), 3rd edition. Online manuscript released August 24, 2025.
- Lucelene Lopes, V Nunes Maria das Graças, Magali S Duran, and Thiago AS Pardo. 2025. A sintaxe no tribunal: apresentando e explorando um corpus jurídico

⁷<https://sites.google.com/icmc.usp.br/poetisa/>

- em português anotado sintaticamente segundo o modelo Universal Dependencies. In *Simpósio Brasileiro de Tecnologia da Informação e da Linguagem Humana (STIL)*, pages 220–232. SBC.
- Lucelene Lopes and Thiago Pardo. 2024. Towards Portparser-a highly accurate parsing system for Brazilian Portuguese following the Universal Dependencies framework. In *Proceedings of the 16th International Conference on Computational Processing of Portuguese*, pages 401–410.
- 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 Daniel Zeman. 2016. [Universal Dependencies v1: A Multilingual Treebank Collection](#). In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 1659–1666, Portorož, Slovenia. European Language Resources Association (ELRA).
- Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Jan Hajič, Christopher D. Manning, Sampo Pyysalo, Sebastian Schuster, Francis Tyers, and Daniel Zeman. 2020. [Universal Dependencies v2: An Evergrowing Multilingual Treebank Collection](#). In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 4034–4043, Marseille, France. European Language Resources Association.
- Adriana S Pagano, Magali Sanches Duran, and Thiago Alexandre Salgueiro Pardo. 2023. Enhanced dependencies para o português brasileiro. In *Proceedings of the 2nd Edition of the Universal Dependencies Brazilian Festival*, pages 461–470.
- Sebastian Schuster and Christopher D. Manning. 2016. [Enhanced English Universal Dependencies: An Improved Representation for Natural Language Understanding Tasks](#). In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 2371–2378, Portorož, Slovenia. European Language Resources Association (ELRA).
- Tianze Shi and Lillian Lee. 2021. [TGIF: Tree-Graph Integrated-Format Parser for Enhanced UD with Two-Stage Generic- to Individual-Language Finetuning](#). In *Proceedings of the 17th International Conference on Parsing Technologies and the IWPT 2021 Shared Task on Parsing into Enhanced Universal Dependencies (IWPT 2021)*, pages 213–224, Online. Association for Computational Linguistics.
- Kaiqiang Song, Logan Lebanoff, Qipeng Guo, Xipeng Qiu, X. Xue, Chen Li, Dong Yu, and Fei Liu. 2019. [Joint Parsing and Generation for Abstractive Summarization](#). *ArXiv*, abs/1911.10389.
- Shuangzhi Wu, Dongdong Zhang, Zhirui Zhang, Nan Yang, Mu Li, and Ming Zhou. 2018. [Dependency-to-Dependency Neural Machine Translation](#). *IEEE/ACM Trans. Audio, Speech and Lang. Proc.*, 26(11):2132–2141.