CONTENTS

Articles

Revamping the SLTev Tool for Evaluation of Spoken Language Translation
Michelle Elizabeth, Ondřej Bojar 5

Let’s Do It Orderly: A Proposal for a Better Taxonomy of Adverbs in Universal Dependencies, and Beyond
Flavio Massimiliano Cecchini 15

Instructions for Authors 66
Abstract

This article describes recent improvements of SLTev, a tool for automatic evaluation of machine translation, speech recognition and speech translation systems. The changes include the implementation of the COMET score for evaluation of machine translation and spoken language translation outputs as well as a fix for the problematic delay calculation for repeated words which favoured longer segments. Additionally, the system outputs of the IWSLT 2022 shared task have also been evaluated using SLTev and a comparison study with another speech evaluation toolkit, SimulEval, has been done.

1. Introduction

Spoken Language Translation or SLT is a prominent task in NLP. In the so-called cascaded approach, it involves translation of speech into various languages and combines automatic speech recognition in the source language and then machine translation into the target languages. On the end-to-end approaches, the intermediate transcription in the source language is not explicitly considered. As with any application in NLP, it is necessary to evaluate the results produced with suitable metrics.

SLTev\(^1\) or Spoken Language Translation Evaluation tool (Ansari et al., 2021) performs the evaluation of the outputs of spoken language translations by reporting the quality, latency, and stability of a candidate output based on its time-stamped transcript and reference translation into a target language.

\(^1\)https://github.com/ELITR/SLTev
This article gives an overview of the work done to improve the SLTev tool. Our main contributions include the addition of a new metric, COMET for evaluating machine translation and the fix for delay calculation so that it takes into account the timestamps of repeated tokens. Also, an evaluation of the system outputs of IWSLT 2022 (Anastasopoulos et al., 2022) was done using the SLTev tool, comparing the scores also to the results of SimulEval, the tool used officially for IWSLT 2022.

We first look at the SLTev tool and the metrics it uses for evaluation. Then we discuss the issues and challenges existing in the current implementation of the tool and the work done to mitigate these issues and improve the tool. Finally, we evaluate the IWSLT 2022 outputs using SLTev.

2. Background

The SLTev tool is an open source tool for evaluating SLT outputs against reference translations and time-stamped source transcripts. It was developed as part of the European Live Translator (ELITR) project (Franceschini et al., 2020) and provided three metrics namely SacreBLEU (Post, 2018) for measuring quality, Flicker for measuring the stability and Delay for measuring the latency of SLT outputs.

Translation quality is estimated using the SacreBLEU tool applied in three different ways within SLTev. The first one considers all completed segments as a single joint segment and compares it with the reference which is also considered as a single concatenated segment. The second variant uses mwerSegmenter (Matusov et al., 2005) to compare candidate and segmented reference outputs. The final variant relies on time-span quality and divides the whole document into chunks or segments of a fixed duration which are then separately evaluated using BLEU and also averaged for the score of the whole document.

Flicker assesses the amount of intermediate output updates which can distract the user by counting the number of words after the first difference between two consecutive output updates. Flicker is reported in two variants: average revision count per second and normalised revision count which are described by Ansari et al. (2021).

The final measure is delay which measures the difference between the time that a target word was displayed and an estimate of when it should have been displayed given the source transcript. The delay is calculated using two approaches. The first one is proportional which estimates the timing of each source word based on partial segments in the golden transcript. These times are then passed to the words in the reference translation proportionally along the sequence of words. The second approach uses automatic word alignment between the source and reference translations to account for word order differences across languages.
2.1. SLTev vs. SimulEval

Another toolkit that is similar to SLTev and has been developed for evaluating simultaneous translation is SimulEval (Ma et al., 2020). It has been used as the evaluation toolkit for the IWSLT Shared Task since its first edition in 2020.

SimulEval is based on a client-server scheme in which the server sends the source input when requested by the client, receives the translation for evaluation from the client and reports various metrics pertaining to translation quality and latency. The client is composed of an agent and a state. The agent is responsible for executing the system’s policy and the state tracks the necessary information for executing the policy when generating the translation. SLTev on the other hand uses a time-stamped golden transcript in the source language, a reference translation and candidate output in the target language to evaluate the translation quality, latency and stability off-line, i.e. from logs and without running the system again.

SimulEval reports BLEU, TER and METEOR for evaluating translation quality and has adapted Average Proportion, Average Lagging and Differentiable Average Lagging for speech translation. However, it does not support any assessment of output; the evaluated systems are not permitted to their older outputs in any way. SLTev reports stability using the flicker metric along with measuring translation quality and latency using SacreBLEU and delay respectively as described previously. An evaluation of the IWSLT 2022 using SLTev and a comparison of the results with the official SimulEval results is reported later in Section 3.3.

3. Improvements to the Current Implementation of SLTev

This section describes some of the issues that the current implementation of SLTev had and the work done to improve the tool.

3.1. Delay Computation

In the existing implementation of delay computation, there was an issue in how the time-stamps were assigned to repeated words in a segment. This problem has been reported by Amrhein and Haddow (2022). The following example can be used to explain the problem:

P 13.18 O
P 14.18 O horror,
P 15.18 O horror, terror, horror
C 16.18 O horror, horror, horror.

SLTev assigned the time stamp of 14.18 to all occurrences of the word “horror”, i.e. it assigned the token the time-stamp of its first occurrence even though later updates actually discarded some of these occurrences. When translating longer segments easily consisting of multiple sentences, the likelihood of encountering tokens that were
previously seen increases. In such cases, all of these tokens would be assigned the time-stamp of their first occurrence. Hence, this tends to favour longer segments in the translation process.

3.2. COMET

COMET (Rei et al., 2020), which stands for Cross-lingual Optimized Metric for Evaluation of Translation, is a popular neural framework for training multilingual machine translation evaluation models. Typically, COMET models are trained with the objective of predicting quality scores for translations. These scores are usually normalized through a z-score transformation and serve as a valuable metric for ranking translations and systems based on their quality.

The COMET library has several evaluation models and we use the default model Unbabel/wmt22-comet-da (Rei et al., 2022). This model utilizes a reference-based regression methodology and is constructed using the XLM-R framework. It has undergone training on direct assessments from WMT17 to WMT20, offering scores within the 0 to 1 range. A score of 1 indicates a perfect translation.

The Unbabel/wmt22-comet-da model is available on HuggingFace and can be downloaded. A list of dictionaries containing the source, candidate translation and the reference is given as input to the model. It generates scores for each set of source, candidate, reference triplet and also reports an overall system score.

For SLTev, the segmented candidate sentences were concatenated together to form one single segment. The same was done for the source and reference segments in order to generate an overall score for the document. The generated system score which is reported by the model in the range [0, 1] has been scaled to [0, 100] in order to be consistent with the SacreBLEU reporting in SLTev.

One issue that was observed during the implementation was that internet connection was necessary in order to download the model to the local system. Currently, this situation is being handled in a way that does not disrupt the flow of the evaluation by handling the exception where the download has failed and moving on to the next metrics.

3.3. Evaluation of the IWSLT 2022 System Outputs

The SLTev tool was used to run an evaluation of the outputs by the models submitted to the IWSLT 2022 Simultaneous Speech Translation task. The language pair was English to German and the outputs of five systems namely CUNI-KIT (Polák et al., 2022), FBK (Gaido et al., 2022), HW-TSC (Wang et al., 2022), NAIST (Fukuda et al., 2022) and UPV (Iranzo-Sánchez et al., 2022) were evaluated. Each system has produced outputs for three latency regimes — high, medium and low — determined by a maximum latency threshold measured by Average Lagging on the Must-C tst-COMMON set.
<table>
<thead>
<tr>
<th>Model</th>
<th>Delay Without Partials</th>
<th>Delay With Partials</th>
<th>SacreBLEU</th>
<th>COMET</th>
<th>Flicker</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUNI-KIT.high</td>
<td>61.24</td>
<td>62.49</td>
<td>25.59</td>
<td>66.57</td>
<td>0.00</td>
</tr>
<tr>
<td>CUNI-KIT.medium</td>
<td>71.39</td>
<td>71.84</td>
<td>23.32</td>
<td>64.98</td>
<td>0.00</td>
</tr>
<tr>
<td>CUNI-KIT.low</td>
<td>79.72</td>
<td>79.93</td>
<td>17.76</td>
<td>60.01</td>
<td>0.00</td>
</tr>
<tr>
<td>FBK.high</td>
<td>90.25</td>
<td>91.91</td>
<td>18.57</td>
<td>54.12</td>
<td>0.00</td>
</tr>
<tr>
<td>FBK.medium</td>
<td>58.51</td>
<td>60.20</td>
<td>15.85</td>
<td>51.22</td>
<td>0.00</td>
</tr>
<tr>
<td>FBK.low</td>
<td>51.07</td>
<td>52.23</td>
<td>8.62</td>
<td>41.18</td>
<td>0.00</td>
</tr>
<tr>
<td>HW-TSC.high</td>
<td>50.31</td>
<td>50.91</td>
<td>9.88</td>
<td>35.09</td>
<td>0.00</td>
</tr>
<tr>
<td>HW-TSC.medium</td>
<td>60.24</td>
<td>60.85</td>
<td>9.82</td>
<td>35.3</td>
<td>0.00</td>
</tr>
<tr>
<td>HW-TSC.low</td>
<td>65.44</td>
<td>66.1</td>
<td>8.31</td>
<td>33.94</td>
<td>0.01</td>
</tr>
<tr>
<td>NAIST.high</td>
<td>77.82</td>
<td>79.37</td>
<td>9.00</td>
<td>38.33</td>
<td>0.00</td>
</tr>
<tr>
<td>NAIST.medium</td>
<td><strong>27.79</strong></td>
<td><strong>28.73</strong></td>
<td>9.16</td>
<td>38.00</td>
<td>0.00</td>
</tr>
<tr>
<td>NAIST.low</td>
<td>38.15</td>
<td>38.34</td>
<td>7.03</td>
<td>39.75</td>
<td>0.01</td>
</tr>
<tr>
<td>UPV.high</td>
<td>238.31</td>
<td>241.12</td>
<td>22.88</td>
<td>62.29</td>
<td>0.00</td>
</tr>
<tr>
<td>UPV.medium</td>
<td>148.25</td>
<td>150.31</td>
<td>19.49</td>
<td>59.97</td>
<td>0.00</td>
</tr>
<tr>
<td>UPV.low</td>
<td>176.52</td>
<td>179.75</td>
<td>12.81</td>
<td>52.06</td>
<td>0.00</td>
</tr>
</tbody>
</table>

*Table 1. SLTev Evaluation of IWSLT 2022 System Outputs*

<table>
<thead>
<tr>
<th>Latency</th>
<th>High</th>
<th>Medium</th>
<th>Low</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson Correlation</td>
<td>0.89</td>
<td>0.864</td>
<td>0.891</td>
</tr>
</tbody>
</table>

*Table 2. Pearson Correlation of IWSLT 2022 System Outputs with respect to SimulEval and SLTev Results*
Figure 1. Scatterplot showing the BLEU scores reported by SLTev and SimulEval for IWSLT 2022 systems for three latency regimes.
IWSLT 2022 used the SimulEval toolkit (Ma et al., 2020) for evaluating the quality and latency of the submissions. The metrics they used were BLEU for measuring translation quality and average proportion (AP), average lagging (AL) and differentiable average lagging (DAL) for measuring translation latency. Using SLTev, the delay measurements calculated using the partial segments and the one which considers only the completed segments, SacreBLEU, COMET and Flicker have been measured. We can see that the translation quality is highest in each system for the high latency regime except for NAIST which has the best BLEU score for medium latency regime. The CUNI-KIT system in the high latency regime has the best translation quality in terms of both BLEU and COMET. NAIST (medium) has the least average delay whereas UPV systems have the highest delays among all the systems. The metrics are reported in Table 1.

The BLEU scores reported by SimulEval and SLTev in the three latency regimes — high, medium and low — have also been reported as a scatter plot in Figure 1. We can observe that the general trend is that SLTev has scored the systems lower than the scores reported by SimulEval except for CUNI-KIT in the high latency regime. HW-TSC and NAIST have been scored much lower by SLTev than by SimulEval, a difference in the range of approximately 5–10 points. This can also be seen for FBK though not to the extent of HW-TSC and NAIST. CUNI-KIT and UPV have similar scores reported by SLTev and SimulEval. Table 2 reports the Pearson correlation for scores of the systems in the three latency regimes.

4. Conclusion and Future Work

SLTev is a comprehensive tool for evaluating the quality of spoken language translation. We wish it became the standard toolkit with a wide adoption.

The work done reported in this article is just the beginning, there is more room for improvement. The implementation of COMET score can be enhanced further by reporting segment-level scores as well. It would also be beneficial to figure out how to download the COMET model available in HuggingFace locally when installing SLTev and not having to rely on a stable Internet connection to generate the score. The bug fix for delay computation should give more accurate results and will no longer favour longer segments since the time-stamps of repeated tokens are being accurately calculated. Additional metrics relevant to translation can be added including average lagging and chrF3 (Popović, 2015).

The tool can be made more versatile by making it platform independent. Right now, it relies on mwerSegmenter which can only be run on Linux systems. Word-error-rate-based segmentation is thus not preformed for quality evaluation when used on other platforms. The readability and reusability of the code can be improved by using more Pythonic constructions. Also, writing and maintaining unit tests, implementing a proper error handling module and detailed logging are some other ways in which the tool can be made more user and developer friendly.
Bibliography


Iranzo-Sánchez, Javier, Javier Jorge Cano, Alejandro Pérez-González-de Martos, Adrián Giménez Pastor, Gonçal Garcés Díaz-Munio, Pau Baquero-Arnal, Joan Albert Silvestre-Cerdà, Jorge Civera Saiz, Albert Sanchis, and Alfonso Juan. MLLP-VRAIN UPV systems for


Address for correspondence:
Ondřej Bojar
bojar@ufal.mff.cuni.cz
Institute of Formal and Applied Linguistics
Faculty of Mathematics and Physics,
Charles University
Malostranské náměstí 25
118 00 Praha 1, Czech Republic
Abstract

With this paper, we hope to highlight the critical issues and “bad practices” that are currently observed in the representation of adverbs in the annotation framework of Universal Dependencies (where they are ideally identified with the universal part of speech ADV), which themselves more generally mirror the conspicuous lack of systematic definitions of this word class in traditional grammars. This fact, on the one side, hampers a useful and meaningful linguistic description of adverbs in Universal Dependencies’ treebanks and beyond, and, on the other side, has unfortunate consequences for linguistic research, such as too confused, impractical results when, for example, querying a treebank for all ADV-tagged words. Therefore, with the aid of minimal data analysis and numerous examples, this contribution tries to raise awareness about this issue, and proposes a revised, more typologically grounded and general framework for the classification of adverbs in Universal Dependencies, and more broadly advocates a more flexible representation of the interplay between word classes and syntactic functions, through the comprehensive concept of “transposition” or “transfer”. While grounded in the Universal Dependencies formalism, the scope of the discussion of this paper is by no means limited to it, and might be of interest to any practitioner of (computationally-oriented) linguistic annotation.

1. Introduction

Despite the historical and consolidated use in (Western) linguistics of the notion of “adverb” as a part of speech, reaching at least as far as grammatical studies in Ancient Rome, surprisingly this word class still seems to lack a clear and systematic
general definition in literature. On the contrary, at a practical level it often appears as a confusing concoction of different morphosyntactic elements. This might sometimes even be pointed out in general works, as by Schachter and Shopen (2007, p. 19f):

“The label adverb is often applied to several different sets of words in a language, sets that do not necessarily have as much in common with one another, either notionally or grammatically.”

or also by Anderson et al. (2003) s.v. Adverb (Grammar) in the *su. Glossary of Linguistic Terms* (which is referred to by the Universal Dependencies’ guidelines):

“An adverb is a lexical category whose members have the same syntactic distribution [...].”

followed immediately later, contradictorily, by:

“The general class adverb is a mixture of very different kinds of words, which cover a wide range of semantic concepts and whose syntactic distribution is disparate.”

On this note, even in the introduction to a comprehensive treatment of the more general concept of adverbial construction such as van der Auwera and Ó Baoill (1998), we find the following “statement of disengagement” (van der Auwera, 1998, §3):

“The notion of adverb or adverbial has not figured prominently in discussions of linguistic typology. This has at least three reasons. First, the category itself is elusive: it is not clear what the defining or prototypical features of adverbs and adverbials are [...] [The book] does not attempt to throw any new light on the definition of adverb and adverbial, but merely presupposes a classical understanding of «adverb» as the word-level adverbial expression, and of «adverbial» as a syntactically optional modifier of primarily nonnominal constituents.”

And so it might well be the case that some works primarily dealing with adverbs do not even define them explicitly, cf. Cinque (1999).

This situation is most evident in language-specific traditions, where, even though this fact is often highlighted,² all these heterogeneous elements are then still kept together, as if by a sort of rationalisation for want of alternatives. Others still, apparently mostly with respect to non-European languages, might eschew the topic completely.³

We claim that this state of disarray is reflected by current annotation practices in Universal Dependencies, too, where the problem becomes acute because of the desire for a typologically valid formalism. We recall that Universal Dependencies

---

²As e.g. for German in Barz et al. (2009, §837): “Adverbs represent a heterogeneous and therefore difficultly definable word class” (translated from German by the author).

³For example, in the comprehensive presentation of Oceanic languages in Lynch et al. (2011), hardly any of the 43 grammar sketches ventures into describing adverbs as a standalone category, which are only shortly mentioned marginally to other “lexical morpheme classes” when portraying Proto-Oceanic (p. 63: “There may also have been a small class of underived adverbs”).
“is at the same time a framework for crosslinguistically consistent morphosyntactic annotation, an open community effort to create morphosyntactically annotated corpora for many languages” (de Marneffe et al., 2021, §1)

and, at the time of writing, its latest release v2.13 (November 2023) contains 259 treebanks for 148 languages and 31 language families.

In the following we propose, if not the ultimate definition of what an adverb is, at least an overview of the problem and directions for an improved annotation of this word class (or the corresponding syntactic function, as it will be discussed) in Universal Dependencies. Section 2 introduces the Universal Dependencies (henceforth also UD) formalism for readers who might not be so familiar with it, as it is the main setting of this paper. Section 3 describes the problem generally and more specifically in UD; Section 4 brings the perspective of this work into focus; Section 5 puts forward a gradient of increasingly “innovative” solutions by also, inevitably, going beyond merely adverbs; Section 6 concludes the paper.

A note about data  The reference for the cited data is the latest release of the UD project, v2.13, published on November 15th, 2023 (Zeman et al., 2023). The languages and treebanks thereof considered in this paper are, ordered by their number of syntactic words:

- German HDT (Borges Völker et al., 2019; 3 455 580 words)
- Latin, all five treebanks, i.e.
  - IT-TB (Cecchini et al., 2018; Passarotti, 2019; 450 517 words)
  - LLCT (Cecchini et al., 2020a; 242 411 words)
  - PROIEL (cf. Eckhoff et al., 2018; 205 566 words);
  - UDante (Cecchini et al., 2020b; 55 519 words)
  - Perseus (cf. Bamman and Crane, 2011, and also Gamba and Zeman, 2023; 29 221 words)
- English EWT (Silveira et al., 2014; 254 862 words)
- Bulgarian BTB (cf. Osenova and Simov, 2004; 156 149 words)
- Greek (modern) GDT (Prokopidis and Papageorgiou, 2017; 63 441 words)

A further set of five other languages, less or possibly not at all (marked with *) familiar to the author, is selected randomly with the aim of widening the typological perspective, especially in §5:

- Armenian* (western) ArmTD (Yavrumyan, 2019; 122 907 words)
- Faroese FarPa (40 484 words)
- Thai* PUD (cf. Zeman et al., 2017; 22 322 words)
- K’iche’* IU (Tyers and Henderson, 2021; 10 013 words)

\[^4\]http://hdl.handle.net/11234/1-5287.
Amharic: att (Seyoum et al., 2018; 10010 words)

For all details, the respective documentation pages, reachable via the main UD website, can be referred to. The selected treebanks all currently conform to the validation standards of UD or are considered to be “legacy” with minor errors, mostly related to recent amendments in the guidelines. When more than one treebank is present for a language, the largest one is chosen. All examples in the text refer to the sentence identifiers (sent_id comment line) in the CoNLL-U files. For all tags of parts of speech and dependency relations, the reference are the respective entries in the UD guidelines, which we sometimes cite in this paper. In the examples, the focused element is highlighted in bold, while other elements of interest are underlined; forms are in italics, while lemmas are in small capitals. All words are transcribed using recognised standards.

All code and data used for this work (especially in §3.2, 4.4, and 4.4.2) are made available in a public GitHub repository, so as to make our results reproducible, and to allow for updates with respect to future UD releases.

2. Overview of the Universal Dependencies formalism

Even if the issue discussed in this paper is of general linguistic interest, it is true that some points (especially in §3.1, 3.2, and §4.4.2) and the concrete design of the proposal of a new taxonomy (as in §5) are particularly rooted in the formalism of the Universal Dependencies project. It is thus in order to give a short introduction to its basic principles here, highlighting those that are of most interest to this work, and to supply a brief explanation of some of the UD labels which are most relevant to the present discussion. Apart from the documentation that can be found on the project’s website, the main cited introductory papers are Nivre et al. (2020) and de Marneffe et al. (2021). Moreover, in Table 1 we give a synopsis and a brief comment on UD labels, focusing on those most mentioned and used throughout this article: there, the acronym POS stands for “part of speech”, the abbreviation deprel for “(syntactic) dependency relation”, while feature is short for “morpholexical feature”.

At its core, UD is “a project that is developing cross-linguistically consistent treebank annotation for many languages” (Nivre et al., 2020, §1)

5https://universaldependencies.org/
6http://quest.ms.mff.cuni.cz/udvalidator/cgi-bin/unidep/validation-report.pl
7https://universaldependencies.org/changes.html.
8https://universaldependencies.org/guidelines.html
9https://github.com/Stormur/OrderlyAdverbs
With regard to syntax, it makes use of a dependency-based paradigm with a lexicalist approach, taking (syntactic) words as its base units (so words depend on each other, and syntactic trees do not include nodes for phrases, collocations, or similar) and giving content words primacy over function words. All of this can be exemplified by the simple sentence

(1)  *And can anyone use military pressure without proof?*

which is analysed as follows:

Here we observe that:
– the sentence is headed by the predicate, here realised by the verbal phrase *can use*;
– in this English treebank, the modal verb form *can* is considered an auxiliary, and so inside this phrase depends on *use* as *aux* (‘auxiliary’), making *use* the head of the whole sentence, or *root*;
– the co-ordinating conjunction *And* introducing the sentence depends on the predication, as *cc* (‘co-ordinating conjunction’);
– in the phrase *military pressure*, the modifier *military* depends on the modified noun as *amod* (‘adjectival modifier’), and thus *pressure* is the head depending on the predicate as *obj* (‘object’);
– in the phrase *without proof*, it is the preposition (function word) *without* that depends on the noun (content word) *proof* as *case* (‘case marking’), and thus *proof* depends on the predicate as head of this *obl* (‘oblique’) argument;¹⁰
– finally, we note that punctuation is assigned a purely formal “relation” *punct*, attached so as not to create non-projective (i.e. discontinuous) configurations, usually to the head of the phrase it appears the nearest, here the *root* *use*.

¹⁰The dichotomy between *core* (basically *A, P and S* in the terminology described e. g. by Haspelmath, 2011) and *oblique* arguments (all the others), as opposed for example to that between (obligatory) *complements* and (facultative) *adjuncts*, is another basic tenet of UD. We are not dealing with this complex topic here, but point to Andrews (2007) as one of the main introductory works on the subject.
Some of these annotation choices contrast with other widespread formalisms, such as those found in the Prague Dependency Treebank\(^\text{11}\) Hajič (1998); Hajič et al. (2017, 2020), or those assumed in Osborne (2019), where e.g. adpositions and auxiliaries head their respective phrases.

A complete annotation in ud considers, besides syntactic relations, also a lexicosemantic layer with lemmatisation, a morphosyntactic layer with part-of-speech tagging, and a morphological layer with the assignment of morpholexical traits in feature-value pairs. So, in our sample sentence we will have e.g. the lemma and for the form And (case-normalised), the part of speech ADP ‘adposition’ for without, a morphological feature Number with value Sing (‘singular’) for proof\(^\text{12}\) and a lexical one PronType (‘pronominal type’\(^\text{13}\)) with value Ind (‘indefinite’) for anyone.

ud’s system acts according to two important distinctions, defining two kinds of “main axes” along which parts of speech (cf. Tables 1, 2 and paragraph below), but also dependency relations, can be laid out. The first one is a “typology of phrasal units” (de Marneffe et al., 2021, §2.1.2) identifying “nominals”, “clauses” and “modifiers”: so a nominal is an element which refers to an entity and is expected to appear as the argument (subject, object, oblique…) of a predicate, and, in terms of parts of speech, we can think of nouns (NOUN in ud) as the prototypical nominals, with the corresponding relations nsubj ‘nominal subject’, obj ‘object’, and so on; similarly verbs (VERB) will be the prototypical roots of clauses, with relations such as csubj ‘clausal subject’ and ccomp ‘clausal complement’, and adjectives (ADJ) the prototypical modifiers, depending as amod ‘adjectival modifier’. The second distinction regards grammaticalisation (de Marneffe et al., 2021, §2.2.1), which implies a cline between two poles. On the one pole, we have autosemantic or content words,\(^\text{14}\) i.e. words with an independent meaning. On the other pole, we find synsemantic, grammatical or function words, i.e. elements which do not have an independent meaning and often serve as grammatical markers specifying the syntactic roles or grammatical categories of other elements, and, if this is the case, in ud’s system they always depend on those elements by means of special relations. Some word classes are easily recognised as tending to one of the two poles: so (proper) nouns (NOUN and PROPN), full verbs (VERB)

\(^{11}\)For the latest version, refer to Hajič et al. (2020).

\(^{12}\)We note that it might be argued that English actually does not morphologically express a singular number in nouns, as opposed to a marked plural, e.g. house-∅ vs. house-s, but this English treebank chooses to annotate it, probably recognising a paradigmatic contrast.

\(^{13}\)Actually extending also to other parts of speech than just PRON ‘pronoun’.

\(^{14}\)Sometimes also “lexical words”, but we will avoid this terminology in this sense, as it is too loaded with different meanings, and thus ambiguous; cf. Haspelmath (2024). Here, in particular, it overlaps with the notion of lexical word as opposed to e.g. symbols or punctuation marks, or of lexical features as opposed to morphological ones.
<table>
<thead>
<tr>
<th>Label</th>
<th>Description</th>
<th>Layer</th>
<th>Grammaticalisation</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADJ</td>
<td>adjectives</td>
<td>POS</td>
<td>autosemantic</td>
<td>The prototypical modifiers</td>
</tr>
<tr>
<td>ADP</td>
<td>adpositions</td>
<td>POS</td>
<td>synsemantic</td>
<td>Comprising prepositions, postpositions and circumpositions</td>
</tr>
<tr>
<td>ADV</td>
<td>adverbs</td>
<td>POS</td>
<td>underdefined</td>
<td>The main topic of this paper</td>
</tr>
<tr>
<td>CCONJ</td>
<td>co-ordinating conjunctions</td>
<td>POS</td>
<td>synsemantic</td>
<td>Distinct from subordinating conjunctions SCONJ</td>
</tr>
<tr>
<td>DET</td>
<td>determiners</td>
<td>POS</td>
<td>synsemantic</td>
<td>Synsemantic counterparts of adjectives, often deictic</td>
</tr>
<tr>
<td>NUM</td>
<td>numerals</td>
<td>POS</td>
<td>synsemantic</td>
<td>In use substantially coinciding with “cardinal number”, and as such effectively a subclass of DET</td>
</tr>
<tr>
<td>PART</td>
<td>particles</td>
<td>POS</td>
<td>synsemantic</td>
<td>Residual function words; underdefined category</td>
</tr>
<tr>
<td>PRON</td>
<td>pronouns</td>
<td>POS</td>
<td>synsemantic</td>
<td>Synsemantic counterpart of nouns, referring deictically to entities instead of naming them</td>
</tr>
<tr>
<td>SCONJ</td>
<td>subordinating conjunctions</td>
<td>POS</td>
<td>synsemantic</td>
<td>Distinct from co-ordinating conjunctions CCONJ</td>
</tr>
<tr>
<td>advcl</td>
<td>adverbial clause modifier</td>
<td>deprel</td>
<td>autosemantic</td>
<td>Clause acting as a non-core modifier of the predicate; clause equivalent of advmod</td>
</tr>
<tr>
<td>advmod</td>
<td>adverbial modifier</td>
<td>deprel</td>
<td>underdefined</td>
<td>Formally required for ADV; see §3.1</td>
</tr>
<tr>
<td>advmod:emph</td>
<td>emphasising word</td>
<td>deprel</td>
<td>synsemantic</td>
<td>Subtype of advmod used by some treebanks; see §5.1.2</td>
</tr>
<tr>
<td>aux</td>
<td>auxiliary</td>
<td>deprel</td>
<td>synsemantic</td>
<td>Required for grammatical verbal elements in periphrastic verbal predicates</td>
</tr>
<tr>
<td>case</td>
<td>case marking</td>
<td>deprel</td>
<td>synsemantic</td>
<td>Required by adpositions associated to nominals, which are equated to morphological case markers</td>
</tr>
<tr>
<td>cc</td>
<td>co-ordinating conjunction</td>
<td>deprel</td>
<td>synsemantic</td>
<td>Required by co-ordinating conjunctions</td>
</tr>
<tr>
<td>det</td>
<td>determiner</td>
<td>deprel</td>
<td>synsemantic</td>
<td>Required by modifying determiners</td>
</tr>
<tr>
<td>discourse</td>
<td>discourse element</td>
<td>deprel</td>
<td>underdefined</td>
<td>Relation of pragmatic nature</td>
</tr>
<tr>
<td>mark</td>
<td>marker</td>
<td>deprel</td>
<td>synsemantic</td>
<td>Required by subordinating conjunctions and adpositions introducing an argument clause</td>
</tr>
<tr>
<td>obl</td>
<td>oblique nominal</td>
<td>deprel</td>
<td>autosemantic</td>
<td>Any argument of the predicate which is not the subject or an object</td>
</tr>
<tr>
<td>PronType=Dem</td>
<td>demonstrative pronominal type</td>
<td>feature</td>
<td>lexical</td>
<td>Element with deictic function, e.g. that, this</td>
</tr>
<tr>
<td>PronType=Int</td>
<td>interrogative pronominal type</td>
<td>feature</td>
<td>lexical</td>
<td>Element appearing in questions, e.g. what, who</td>
</tr>
<tr>
<td>PronType=Rel</td>
<td>relative pronominal type</td>
<td>feature</td>
<td>lexical</td>
<td>Element with anaphoric function appearing in so-called relative clauses, e.g. which, whose</td>
</tr>
<tr>
<td>VerbForm</td>
<td>form of verb or deverbative</td>
<td>feature</td>
<td>morphological</td>
<td>“External”, non-verbal syntactic role of a verb form still acting as a predicate</td>
</tr>
</tbody>
</table>

*Table 1. Synopsis and brief explanation of the main UD labels used throughout this paper.*
and adjectives (ADJ) are instances of content words; conversely, auxiliaries (AUX), adpositions (ADP), conjunctions (CCONJ and SCONJ) and generic particles (PART) supply examples of prototypical function words. While this distinction is clear in much of the literature about UD, as in de Marneffe et al. (2021) in primis, there remain some word classes whose status of grammaticalisation is more controversial: determiners (DET), numerals (NUM), and pronouns (PRON). In the framework of this work, with good reasons but possibly against the general consensus (especially for what concerns pronouns, or pro-forms in general), we consider all these three classes to have a synsemantic nature, as discussed in the following paragraph. This is especially reflected in Tables 1 and 2, and in Section 3.2.

**Grammaticalisation status of DET, NUM and PRON in UD and in this paper** The main issue with the three UD word classes DET, NUM and PRON is that they can occur (though at different degrees across languages) in the same contexts as NOUNs, i.e. as heads of arguments of the predicate. While this might favour the identification of their members as autosemantic, there are indeed some fundamental reasons to do otherwise. One is internal to UD’s system as defined in the guidelines: DET and NUM have dedicated dependency relations, respectively det and nummod. The former is classified under “Function Words”;¹⁵ for NUM it is said:

“A numeral is a word, functioning most typically as a determiner, adjective or pronoun [...] Note that cardinal numerals are covered by NUM whether they are used as determiners or not [...] Other words functioning as determiners (including quantifiers such as many and few) are tagged DET.”

This goes to show how both DET and NUM are understood as synsemantic word classes in UD’s system, and moreover NUM as a subclass of DET. However, a different treatment seems to be in place for the class PRON, judging e.g. from de Marneffe et al. (2021, §2.2.1):

“We expect the words that enter into the main syntactic relations to be autosemantic [...] typically verbs, nouns, adjectives, or adverbs, as well as corresponding pro-forms with a contextually determined referential meaning.”

But this is at odds with the classification of determiners and numerals as synsemantic: in fact, we argue that their being “contextually determined” is an aspect of synsemanticity, which is common to determiners and pronouns (and pro-forms in general): literally, not possessing an independent meaning, be this from a syntactic (ADP, AUX, CCONJ, SCONJ, PART) or a semantic (DET, NUM, PRON) point of view. In this paper, we will use this broader, at least with respect to de Marneffe et al. (2021), notion of

¹⁵[https://universaldependencies.org/u/dep/index.html](https://universaldependencies.org/u/dep/index.html)
synsemanicity. We notice that this stance is apparently supported by passages of UD’s
guidelines such as (about PRONs):

“[i]t is not always crystal clear where pronouns end and determiners start.”

Moreover, words of the PRON class can depend on others with the synsemantic rela-
tion det. In fact, we do expect word classes of the same type of “phrasal unit” to have
very similar if not identical distributions, as it happens for ADJs and DETs (usually
conflated by traditional grammars). In any case, we contend that the main point of
the present work is not affected by sharing or not this position: the mixed nature of
the ADV class will be made clear in Section 3.2.

If, then, DET, NUM and PRON are acknowledged as having a synsemantic nature, with
regard to parts of speech the two main axes of type of phrasal unit and grammatical-
isation concur to define three main “dyads” in UD’s system: NOUN vs. PRON for nomi-
nals, VERB vs. AUX for clauses, and ADJ vs. DET for modifiers. Temporarily disregarding
ADV, the other 10 labels in UD’s part-of-speech system represent either subcategories
of the main ones (PROPN ‘proper nouns’ with respect to NOUN, and NUM ‘numerals’
with respect to DET), purely grammatical elements belonging to no particular type
of phrasal unit (ADP ‘adpositions’, CCONJ ‘co-ordinating conjunctions’, SCONJ subor-
dinating conjunctions, PART ‘particles’), non-categorisable elements (INTJ ‘interjec-
tions’ and X ‘other’), and non-lexical elements (PUNCT ‘punctuation marks’ and
SYM ‘symbols’). In the rest of this work, we argue that the position of the category ADV
‘adverbs’ and the relation advmod ‘adverbial modifier’ inside UD’s schema is under-
defined with respect to many layers, in a way that undermines the soundness of the
whole system, so that some reorganisation needs to take place.

3. The “adverb” issue: from the beginnings to Universal Dependencies

The notion of “adverb”, though being a staple of part-of-speech tagsets, remains
problematic when it comes to pinning down a precise and useful definition of it, par-
ticularly in view of (automated) linguistic annotation. Also, the problem encom-
passes and interweaves at the same time the definitions of parts of speech and de-
pendency relations, as will be examined in Section 4. All of this makes it difficult
to even find truly comprehensive examinations of this word class in literature. We
do find some, like Ramat and Ricca (1994), whose endeavour is partly shared by the
present work, and where the definition of “prototypical adverb” closely resembles
our class of “core” adverbs (§4.3): but even in this case, we claim that the levels of
form and function are (possibly due to a lack of an annotation formalism of reference,

---

16 We further notice that other sources describing UD’s system, such as Zeman (2018, §3.4, Table 3.15), are
also very explicit as to the synsemantic position of PRONs as opposed to NOUNs.

17 We refer to Zeman (2018, §3) for an overview, where a tag for “adverb” always appears, starting from
RB in the Penn Treebank’s tagset (Marcus et al., 1993, §2.2).
as here, ultimately mixed as in the kind of contextual annotation that we criticise (cf. §4.1), in a sense projecting syntactic function back to the word class of any (lexical) element realising it. This is the only way we see how, e.g. in Latin, *merito* ‘having deserved (it)’, i.e. ‘with merit’ can be considered an adverb, while in “*maximo merito* ‘with the highest merit’ there is no adverbial lexeme!” (Ramat and Ricca, 1994, §1, p. 290). Here, the derivational status of some elements is mentioned, but finally disregarded: a form like *merito* can only be labelled as invariable if it is arbitrarily detached from the paradigm it belongs to (in this case that of the verb *mereo* ‘to deserve’; cf. §4.4.2). Further, in that work the focus seems rather to lie on a semantic classification of adverbial elements, with which we are not directly concerned here.

Trying to go to the roots of the issue at hand, since in Western linguistics the term *adverb* itself (as a calque from Greek *epírrēma*, literally ‘(what is) beside the verb’) and its theoretical definition originate from the Ancient Roman grammatical tradition, it can help our understanding of the current situation to look at how adverbs are treated in modern literature about Latin. Latin is known for an extensive grammatical tradition, significantly influenced by the earlier Greek one (cf. Law, 2003, §4.3ff), and itself having influenced most other Western grammatical traditions, which themselves nowadays contribute to shaping *ud*. This has as a consequence that what we find in dedicated literature about Latin is also mirrored by the grammatical literature of many other languages. One can thus read there as a constant about adverbs:

“Under the label «adverb>, morphologically and **syntactically** very different terms are stowed” (Marouzeau, 1949, p. 11)

“absence of formulation of criteria” (Pinkster, 1972, p. 33)

“[…] building up a very heterogeneous set, formally as well as functionally” (Ricca, 2010, p. 109)

“Adverbs constitute a very heterogeneous word class, both in respect to their forms and in respect to their meanings and syntactic functions” (Pinkster, 2015, p. 65)

This vagueness is exemplarily reflected in the *ud* guidelines for the part of speech *ADV*:

“Adverbs are words that typically modify verbs for such categories as time, place, direction or manner. They may also modify adjectives and other adverbs, as in very briefly or arguably wrong.”

Which is more or less equivalent to what we read in papyri of the first century AD:

---

18 On the topic see e.g. Pinkster (1972, §3)

19 All cited sources are standard works about adverbs in Latin, or Latin in general, and will be cited again in this work; Bos (1967); Van Laer and Fruyt (2009) can also be added to the list for further reference. Boldfaces in quotations are by the author.

20 Translated from French by the author.
“The adverb is an uninflected word form which is placed before or after the
verb and not compounded with it, signifying quantity, quality, time, place,
negation, agreement, prohibition, exhortation, interrogation, exclamation,
comparison or doubt.” (cf. Law, 2003, p. 57)

Problems arise when we inspect more closely the elements given as examples in UD’s
guidelines: one immediately notices that some which are usually tagged as ADV in
the treebanks, such as up, where and others, can hardly be said to modify adjectives
(?an up good book, ?a where good book) or adverbs,21 as opposed to exactly (or any other
deadjectival adverb in -ly; cf. §4.4), which we can observe as the modifier (see §4.1 for
a definition) of an adjective (ADJ) in a copula in

(2)  I’m not exactly certain
    – English-EWT email-enronsent35_01-0081

but also of a head noun (NOUN) in a noun (here, prepositional) phrase in 22

(3)  I’m going to go at exactly the same time
    – English EWT email-enronsent34_01-0007.

This latter possibility is not mentioned in the guidelines, which later even feel the
need to single out a specific type of adverb, “pronominal adverbs” (§5.1.4).

This very brief overview shows that the allegedly unitary ADV class is actually
forced to contain (morpho)syntactically very different elements, analogously to what
is reported for Latin. This contrasts, for example, with UD’s definition of another class
of modifiers, ADJ (“Adjectives are words that typically modify nouns and specify their
properties or attributes.”), whose behaviour is decidedly more uniform, whereas vari-
ation rather takes place at a semantic level23 (with which UD is not really concerned).
We thus advocate the need for a rethinking of ADV in UD, which possibly involves other
parts of speech, too, in order to achieve a more coherent, linguistically representative
annotation. We now proceed to highlight more UD-specific critical points of the cur-
rent ADV class.

21 In fact, in no English UD treebank (as of v2.13) words with lemma WHERE or UP are annotated as modi-
ifying (i.e. with dependency relation advmod, see §3.1) an adjective (ADJ), or an adverb (ADV) which is not
part of a predicate, or is itself a so-called “pronominal adverb” (here, there, front; see below and §5.1.4), with
the single exception of close (in the phrase get up close in English-EWT answers-20111108094831An0jgr_ans-
0081), which however oscillates between adjectival and adverbial status.

22 The fact that the modified element here is time and not the predicate am going to go is unambiguously
signalled by the position of exactly inside the prepositional phrase, i.e. after at, and outside the noun phrase,
 i.e. before the determiner the.

23 For example, adjectives are known to distribute over a semantic hierarchy, from more “noulike” to
more “verblike”, which can govern among other things the expression of degree or the use of a copula
3.1. Syntactic relations and part of speech of adverbs in UD: redundancy

Formally, even if not completely explicitly, UD requires ADVs in their modifying function to be paired with the advmod dependency relation; conversely, advmod should only be assigned to words tagged as ADV. This is substantiated by the warning “«advmod» should be «ADV» but it is _” issued by the UD validator script. Nonetheless, in the same script we read that also parts of speech ADJ, CCONJ, DET, PART and SYM are admitted. So, on the one side UD enforces redundancy, similarly as for DET & det (part of speech and relation for determiners; see Table 1 in §2), but on the other side it “tolerates” some annotation practices that are apparently common across treebanks, introducing a covert and unexpected behaviour. In general, we argue that a similar redundancy is not desirable, and that combinations of non-ADV words and the advmod dependency relation could be naturally included in UD’s framework: such a possibility would avoid obscuring transparent patterns such as the English ADJ exact producing the ADV exactly by means of the productive -ly suffix, as a consequence of suppressing morphological features not pertinent to the imposed class of ADV in the annotation (§4.4). This alternation might be seen as a rather inconsistent annotation practice, against the UD desideratum of “rapid, consistent annotation” (de Marneffe et al., 2021, §5, point 3). Sections 4.1 and 4.4 deal with this issue in-depth, and with the repercussions of the fixed ADV/advmod coupling on the representation of lexicon.

3.1.1. Subtypes of advmod

Beyond the basic advmod relation, 47 languages in UD define various subtypes of this dependency label. Some appear rather widespread, like emph (27 languages, including Latin with IT-TB, LCT, UDante, and Western Armenian), others more specific, e.g. cau, deg, eval, foc, freq, tfrom, lmp, lto, mmod, shared by a Uralic block (Erzya, Komi Permyak & Zyrian, Moksha, Skolt Sami; cf. Partanen et al., 2018). Some are mostly semantic in nature, e.g. tto ‘till when’ vs. tfrom ‘from when’ (Hungarian), but others point to actual (morpho)syntactic distinctions (cf. classification in §5): it is the case of emph for an “emphasizing word, intensifier” and neg for negations (see §5.1.2). We argue that if some words always select a relation subtype on the grounds of distinctive syntactic behaviour (and not merely semantics), and vice versa this subtyped

---

24 Barring ellipses and co-ordinative constructions.

25 Here and in the following we refer to lines 1475–1483 of the script validate.py, as retrieved from https://github.com/UniversalDependencies/tools (as of January 19th, 2024).

26 Cf. §4.4, and also the issue https://github.com/UniversalDependencies/docs/issues/617.

27 Considering the English ewt corpus, out of 709 unique lemmas of words tagged as ADV, more than half, 381, are derived from adjectival bases (including determiners and participial forms) by means of the -ly suffix.
relation is always realised by such words, then these words actually identify their own (sub)class (i.e. part of speech) and/or dependency relation, at least on some axis (cf. Table 2 and overview in §2).

3.2. Asymmetry of UD’s part-of-speech system and undesirable consequences

Table 2, following de Marneffe et al. (2021, §2.1.2), themselves referring to Croft (1991), shows UD’s part-of-speech system in the framework of this work along its main axes of syntactic functions (or “types of phrasal unit”) and autosemantic/synsemantic dichotomy (de Marneffe et al., 2021, §2.2.1), as discussed in Section 2. Table 2 does not include the classes SYM (symbols) and PUNCT (punctuation marks), as they do not represent linguistic entities; INTJ (interjections) is also excluded, since it lies outside of morphosyntax (cf. Anderson et al., 2003 s.v.), just like X, which is a residual non-class; subclasses are grouped together with the main ones.

Apart from the three main dyads corresponding to the types of phrasal units (corresponding on their autosemantic side to parts of speech NOUN, VERB and ADJ) and a trail of synsemantic parts of speech (ADP, CCONJ/SCONJ, PART) with no defined overarching functions, ADV stands ambiguously in the middle. This class is indeed not specified for synsemanticity or autosemanticity, but actually contains elements of both kinds, cf. VERY and EXACTLY as seen in the beginning of this Section: the former is an intensifier for members of the ADJ class28 with no true autonomous meaning, comparable to what MOST (tagged as ADV for more than half of its occurrences, 102 out of 186, in English EWT) is with respect to Degree=Sup, while the latter can depend on members of the VERB class or predicates in general, and derives its meaning from its autosemantic ADJ base EXACT.29 This goes to show that, as of now, ADV straddles a basic distinction of UD: again, we observe the catchall tendency derived from traditional

28 Apart from a single spurious annotation as root, in

(i) (like VERY not in a good way)
- English EWT answers-20111107194625AAAUBXm_ans-0011

where way ought to be the head, and very acts as a modifier of the whole noun phrase, like not.

29 And this independently from the fact that its scope be only the predicate or the sentence as a whole.
grammars.

This ambiguity can be shown more generally with a small experiment: if, for a given treebank, the two classes of a dyad, say ADJ and DET, are conflated into a single one, which in this case we might call MOD (for “modifier”), we see the synsemantic elements consistently making up the majority in a ranking by frequency. Parallelly, we see this happening for the ADV class, too (in the following lists, synsemantic elements are underlined):

- German HDT

- Latin (it-tb only)

- English EWT
  MOD THE, A, THIS, GOOD, ALL, SOME, ANY, GREAT, OTHER, NO
  ADV SO, JUST, WHEN, VERY, ALSO, HOW, NOW, EVEN, THEN, THERE

- Bulgarian BTB

- Greek (modern) GDT

30 The entries labelled as synsemantic under MOD follow their annotation as DET in their respective treebanks: however, we notice that the elements OTHER, ÓLOS ‘all’, POLYΣ ‘many’, DRUG ‘other’ should also be considered determiners (of contrast and quantity), analogously to ALIUS ‘other’, ALLE ‘all’, ANDERER ‘other’. The judgment about the synsemantic nature of entries under ADV is instead by the author and will be expounded in §5.
Here some patterns emerge: we see parallel synsemantic elements (cf. §4.3) based on the same roots, even in paradigmatic relation, such as POLÝS ‘many’ and POLY ‘a lot’ in Greek (the latter being the neuter singular form of the former). Also, the assortment of the most frequent adverbs is remarkably similar for all languages: one finds the same kinds of elements in many languages, such as demonstratives SIC/SO/SÓ/TAKA, focalisers ETIAM/ALSO/AUCH (see §5.1.2), besides deictic/relational terms which are admittedly more uncertain as to their synsemantic or autosemantic nature (see §5.1.4 and §5.1.5), and which we can tentatively arrange along a gradient:

- clearly synsemantic (i.e. purely deictic, in space or time), such as THEN, THERE;
- mixing synsemantic and autosemantic elements (represented by a dashed line), such as SUPRA ‘above’ (both deictic and with meaning of a specific location);
- clearly autosemantic, such as NOW, SÍMERA ‘today’, VČERA ‘yesterday’.

The few uncontestably autosemantic terms emerging from this conflation are adjectives of nationalities of importance to the respective treebanks, such as DEUTSCH ‘German’, EURÓPAIKÓS ‘European’, BĂLGARSKI ‘Bulgarian'; some universally extremely common qualifying adjectives, such as GOOD, GREAT/MEGÁLOS/GOLJAM (cf. also MAGIS, from MAGNUS ‘great’), NÉOS/NOV ‘new’; and finally text-specific DIUINUS ‘divine’, coming from the theological treatises in the IT-TB, and SÍMANTIKÓS ‘meaningful’. Only farther into the ranking autosemantic elements start appearing more consistently in the class ADV, like the deadjectival ACTUALLY, LONG, LATE in English EWT (from 35th rank onward), or SIMPLICITER ‘simply’ and PERFECTE ‘perfectly’ in Latin IT-TB (from 28th rank onward).

A similar distribution appears to have the side effect of creating unwanted results in a research involving the class ADV in any of these treebanks: most of the retrieved elements will be synsemantic, and they will possibly not correspond to the intuition of what an adverb should look like, if we take a word like exactly as its prototypical member. By the way, this bipartite perspective is partly shared with some dictionaries which are sometimes even used as references for the annotation: so, in Latin, among others, IGITUR ‘therefore’ is not an adverb for e.g. Lewis and Short (1879, s.v.) (cf. §5.1.3), as opposed to SIMPLICITER ‘simply’ (which is traced back to SIMPLEX ‘simple’). It happens so that, in the current state of things, these two subclasses of ADV are unfortunately not separable without resorting to external tools.31 Also, this structural fuzziness between autosemantic and synsemantic elements hinders nucleus-based approaches to parsing, like the one presented by Nivre et al. (2022). In sum, these circumstances do not seem well suited from either perspective: typological representation on the one hand, and linguistic queries at different levels of proficiency on the other hand. This problem is by all means not limited to UD, but, given the imprint of traditional grammars, will be found in many other treebanks, or annotated corpora.

---

31 As the derivational lexicon Word Formation Latin might be for Latin; see Litta and Passarotti (2019) and cf. §4.4.2.
The solution to this undesirable state of affairs is then to refocus the **ADV** class by implementing the same autosemantic/synsemantic split as for the other classes. We notice that this might entail a more or less pronounced restructuring or rethinking of other classes, too (cf. especially §5.1.5); conversely, not every adverb in a traditional sense will be retained as **ADV**.

4. What adverbs are, and what their function is: a working definition

The asymmetry of the **ADV** class inside the **UD** system is a reflex of the historical precedence given by traditional grammars to morphology (and even semantics) over syntax: adverbs are considered indeclinable, isolated elements, and are grouped as such with other function words (see Pinkster, 1972, §4.1). This is for example the stance of the *Index Thomisticus* for Latin, with a single class for all invariable elements, as opposed to a nominal, a verbal and a mixed “participial” one (Busa, 1974–1980). This choice, however, creates internal incoherences as discussed in Section 3. In this Section, we will try to give a definition of the word class “adverb” and of the function associated with it, from which the proposal of a framework for a better annotation in **UD** (and beyond) will ensue.

4.1. Decoupling form from function

An explicitly stated *desideratum* of **UD** is to avoid what might be called “contextual annotation” (de Marneffe et al., 2021, p.262): this is the assignment of any feature, from parts of speech to dependency relations, mechanically depending on the context a word appears in. A classical example from Latin are the *pronoun-adjectives* or *adjective-pronouns* of traditional grammars, like *demonstrative pronouns*, which “are used either adjectively or substantively” (Greenough et al., 2014, §296). In **UD**, this statement translates into tagging an element such as *HIC* ‘this (one)’ as **DET** in a phrase like

(4) **hoc opere**
    ‘(in) this work’
    – Latin *hoc opere* dev-s35

and as **PRON** in another phrase like

(5) **hoc probare**
    ‘to verify this’
    – Latin *hoc probare* dev-s363

---

32 Traditional grammars of Latin do not apply a distinction based on the grammaticalisation cline as **UD**, and thus “adjective” can be translated either as **ADJ** or **DET** in **UD**; *HIC* is **DET** because it is purely deictic (proximal) in function (see Lewis and Short, 1879, s.v.), and carries no other semantic content.
The totally predictable parameter here is merely being head of the argument of a predicate or not, at the cost of losing any information about the distribution of determiners in Latin: for example, that they can head noun phrases more easily than in English, or other languages. The constraint that \texttt{advmod} be realised only by members of \texttt{ADV ($\S 3.1$)} is indeed such a form of part-of-speech–based contextual annotation in \texttt{ud}. However, we observe in \texttt{ud} a tendency to move away from it, e.g. allowing for case/marker alternation for \texttt{ADP}, the possibility for \texttt{PRON} to receive the det relation\(^{33}\) (cf. Höhn, 2021, for an extension to personal pronouns), or the guideline amendment regarding deverbal connectives.\(^{34}\) This is possible because the cornerstone of \texttt{ud} is syntax, and the spread of annotation over four main orthogonal levels (lemmatisation, parts of speech, morpholexical features, dependencies) allows for a rich investigation of typological patterns through their interplay. The first insight re \texttt{ADV} is thus to decouple the syntactic adverbial function (\texttt{advmod}) from the word classes concretely realising it each time. So, there is only to gain in terms of linguistic analysis from observing whether and how members of \texttt{ADJ} (primarily), but also \texttt{DET}, \texttt{NUM}, \texttt{PRON}, \texttt{NOUN}, \texttt{VERB} can or cannot occur as modifiers, possibly assuming a specific adverbial form, at the same time keeping the unitarity of the dependency relation \texttt{advmod}. This perspective does not go as far as to conflating all kinds of modifications (as proposed e.g. by Gerdes et al., 2018), but is more similar to the strategy of an “umbrella” \texttt{Adv} relation as seen in the Prague Dependency Treebank (Hajič, 1998, §2.2,\(^{35}\) and Hajič et al., 2017), as opposed to locking a dependency relation to a specific part of speech.

### 4.2. The adverbial function

If function is decoupled from form, we acknowledge that a word belonging to a given non-\texttt{ADV} part of speech can more or less “easily”, i.e. requiring a more or less explicit specific marking, carry out the function of adverbial modification: this can happen by means of special, dedicated forms which use neither nominal nor verbal strategies. To avoid falling back into contextual analysis (\S 4.1), definitions of parts of speech need to be semantically grounded (at least partly, as argued by Croft, 2001, §2, and for \texttt{ud} by de Marneffe et al., 2021, p. 261). This is actually already the case in \texttt{ud} when e.g. the class \texttt{ADJ} is defined as

“specify[ing] their (sc. of nouns) properties or attributes”.

So, the key feature common to the classes \texttt{VERB} (processes), \texttt{ADJ} (properties) and \texttt{ADV} seems to be relationality: they are all predicative in nature.\(^{36}\) As variously (al-
beit somewhat dishomogeneously) expressed in literature (e.g. Tesnière, 2015, Ch. 32; Hengeveld, 1992, §4.3; Schachter and Shopen, 2007, §1.4), it emerges that the syntactic function of adverbs is that of modifying non-nominal elements, which are often explicitly listed as “verbs, adjectives and other adverbs”. As discussed in Section 3, some elements usually referred to as adverbs do modify nouns or noun phrases, so we argue that a better and more general definition is that the adverbial function is to modify predications, that is, to express properties of a predication or, in other terms, to metapredicate.  

Modification in this sense has to be taken strictly as one type of phrasal unit, as in de Marneffe et al. (2021, §2.1.2), and is different from nominal or verbal strategies. In a broader sense, every argument and every subordinate clause, every single element, “modifies” the sentence it is part of. This vague notion of modification, which seems to be at the basis of the traditional mixed-bag adverb class (cf. §1), is not very useful. In metapredication we explicitly want to include only strategies specifically used by each language for modification: as will be discussed in Section 4.4 and elsewhere, these might be shared with the classes of ADJ/DET, or be peculiar to adverbs. In general, we will consider metapredication mutually exclusive with, but parallel to, nominal oblique arguments (represented by relation obl in UD) and embedded adverbial clauses (represented by relation advcl in UD). Put in other words, we argue that the argument of a predication cannot be considered an adverb; this point of view will be further developed in Section 5. A further important consideration here is that noun phrases, under given conditions, can be considered as (parts of) predications by some elements, e.g. focalisers, and so can receive adverbial modification (§5.1.2). The syntactic function of adverbs, represented by advmod in UD, is thus expressed not in terms of parts of speech, but of relations: without further assumptions, this naturally accounts for possibly infinite chains of modifications, since metapredication is recursive, and an adverb is itself a predication. On the contrary, adjectives cannot modify other words of their class in the sense of a mod, since this kind of modification is directed to referents, i.e. prototypically NOUNs (cf. Table 2). The specific dependency relation tag used in the annotation to represent metapredication can then vary, according to the grammaticalisation status of the adverb, as will be discussed in Section 4.5.

---

37 This term being often vague and itself underdefined, cf. Pinkster (1972, §4.3).

38 Schachter and Shopen (2007, §1.4) speak of “modifiers of constituents other than nouns”, but this still does not seem to take into consideration the matrix predication; for an example regarding Latin, Ricca (2010, §1), too, does not acknowledge nominal modification, while Pinkster (1972) does.
4.3. Autosemantic and synsemantic adverbs: core word classes

The adverbial function as described in Section 4.2 needs a part-of-speech dyad of reference (see Table 2 and §2), representing elements which are primarily (if not exclusively) associated with it (cf. §2). One member of this dyad should be clearly and intuitively ADV, which we propose for autosemantic adverbs; for the other, we put forward a repurposing, or extension, of PART, as of now a rather undefined and officially marginal class in UD’s part-of-speech system. Incidentally, it could be argued that some of the elements already regularly tagged as PART, such as negations or interrogative particles, can indeed be subsumed under the label of synsemantic adverbs (§5.1.2), and in fact they are treated as such by many traditional grammars.39

The resulting dyad ADV/PART of “core” adverbs (also “true” or “primary”, cf. Pinkster, 1972, p. 65) would be mostly composed of elements not regularly or synchronically derived from other bases, and would be more or less open, specifically to each language. So, for example, Latin notably presents words like CLAM ‘secretly’ or PALAM ‘openly’,40 among its exponents of the ADV class, or also SAEPE ‘often’ or EMINUS ‘from a distance’. The latter is clearly related to MANUS ‘hand’, but is not part of its paradigm and has followed its own morphophonetic evolution. On the synsemantic side, we have NAM, a discourse connective, or NE, an interrogative particle (see §4.4.2 for an ampler discussion about Latin). Section 5 will explore more in detail the composition of the class of synsemantic adverbs, drawing distinctions between them and other word classes.

4.4. Core and derived adverbs: strategies and the notion of transposition

Core adverbs ADV/PART are opposed by derived adverbs. While any element expressing metapredication would also regularly receive an adverbial dependency relation (see §4.5), its part of speech does not have to be enforced by its function: contextual annotation (§4.1) is best avoided. In the framework proposed by this paper, we deem that, especially for a multilayered annotation such as that of UD, it is much more fruitful to consider and implement the general phenomenon designated as, among other terms, “transfer” (Tesnière, 2015, P. iii) or “transposition” (Spencer, 2013, §3.3). The fundamental insight here is that inflectional, paradigmatic variations are not limited to single word classes or functions (Haspelmath, 1996): this principle is in fact

39See for example Greenough et al. (2014, §217e) for Latin, where “negative particles” are part of the “classification of adverbs”.

40While these words are related respectively to the same roots of CELO ‘to conceal from’ and PLANUS ‘flat’ with an accusative-like ending (de Vaan, 2008, s. v.), there is no clear nor productive synchronic derivation process producing them, so lexicographically they have to be considered independent; this does happen e.g. in the Word Formation Latin lexicon (Litta and Passarotti, 2019), which however applies a rather restricted notion of derivation/inflection (cf. §4.4.2).
already accepted in UD when participial forms, i.e. verb forms with the function of
attributive predications, more commonly labelled as “verbal adjectives” (cf. Cecchini,
2021 re UD), are still usually and foremostly labelled with VERB, and not with ADJ (a
choice which would exemplify again contextual annotation depending on syntactic
function, cf. §4.1).

Some treebanks appear to pursue this strategy already to some extent. In Table 3,
for each well-represented (>5% of total occurrences) part of speech taking advmod in
the treebanks of our sample, the most frequent (>5% of total occurrences) lemmas are
shown (if any: in some cases, lemmas are too sparse to reach this threshold). It turns
out that lemmas so detected almost exclusively pertain to function words; notably,
PARTs are overwhelmingly negations (cf. §3.2). However, contrary to this trend, in
the German HDT as much as 20% of words receiving an advmod relation are tagged as
ADJ. This goes hand in hand with the known fact commonly stated as German “con-

Table 3. Distributions of parts of speech over all words receiving the dependency relation
advmod, with any subtype, per treebank in our sample.
verting” adjectives directly into adverbs (Barz et al., 2009, §478), a feature which possibly sets it apart from the “Standard Average European” type it is in the core of (cf. Haspelmath, 2001). Other treebanks look more mixed and there probably is some noise in data, but advmod-taking members of ADJ are seen to some extent in Latin proiel (4.9%) and Thai, too, alongside CCONJs in Latin llct and NOUNs in (Amharic). In most other treebanks, similar regular, though marked adverbialisation strategies (such as -(i)ter/-e derivation in Latin, Ricca, 2010, §2.2.1: par → pariter ‘equally’; -ōs in Greek: synepíς → synepós ‘consequently’; etc.) are obscured by redundant adverbial annotation (§3.1). It is still possible, nevertheless, to let some of these strategies emerge from data by comparing forms of ADV- and ADJ-tagged words.

The most prevalent detected strategies are shown in Table 4, where also the corresponding strictly morphological features of the ADJ forms are summarised (comma notation points to alternation of the values in annotation, parentheses to the values not being consistently annotated). We notice a remarkable uniformity: while bare-form adjectives in adverbial function look universal, many Indo-European languages also share the adoption of neuter, singular or plural, forms, and this ostensibly throughout all their history (cf. Ramat, 2009; Hummel, 2013, for Latin, but also more generally). German, as already noticed, looks synchronically more committed to the bare-form strategy than other European languages, but typologically it is perfectly aligned to a universal tendency which we systematically observe in other languages, too. The neuter strategy is also well-attested in Latin, where it is especially common for adverbs formed from determiners (also cf. §4.4.2), such as quantum ‘as many/much’ from quantus ‘how many/much’, tantum ‘so much/only’ from quantus ‘how many/much’, tantum ‘so much/only’ from quantus ‘how many/much’.

---

41 However, in some contexts one might be tempted to interpret this as secondary predication (the difference being the choice of advcl instead of advmod; see s.v. advcl : pred in ud guidelines). Be that as it may, of course German, too, has a class of core adverbs, mostly represented by the tag ADV, such as gern/gerne ‘gladly’, with no synchronous derivation (Pfeifer et al., 1993, s.v.); cf. footnote 40 for Latin.

42 On the contrary, the 100% for Faroese probably arises from the automatic advmod ⇒ ADV conversion rule.

43 Many of which are confirmed by literature and descriptive grammars; apart from German, here we can point e.g. to Scatton (1993, p. 208f) for Bulgarian: “Certain adjectival forms function productively as adverbs: (1) masculine singular (or plural) forms of adjectives […] (2) neuter singular forms of qualitative adjectives […] (3) colloquially, definite feminine singular forms”; or to Polian (2017, §3.7) with regard to the Mayan language K’iche’: “Many words are used adverbially in Mayan languages, some of them exclusively so, but generally without a specifically adverbial morphology. Adjectives commonly develop adverbial uses (e.g. ‘good’ for ‘well’), and some nouns and verb forms also do.”.

44 That is, excluding features of the lexical kind, such as Style or NumType.

45 We can guess this for Faroese forms, too, knowing from dictionaries that the citation form of langt is langur ‘long’ (similarly as for the related Icelandic and Norwegian), but for some reason Gender and Number are not annotated for ADJs in the treebank.
Table 4. Detection of major deadjectival strategies for adverbialisation arising from data in the languages in our sample.

TANTUS ‘that many/much’, or quid ‘why/how/what way [comes that...]’ from quis ‘who/which one’, as in

(6) **quid** propugnaculis et pinnis urbem arnasse iuvabit

‘What shall it avail you […] to have fortified yourselves with bulwarks and battlements’
– Latin UDante Epi-93

where quid cannot be interpreted as an object of iuvabit ‘it will aid’. In Latin, the neuter strategy is also the only possible one for synthetic comparative forms in the adverbial function.46

The practice of enforcing the part of speech ADV for words in adverbial function is surely influenced by traditional grammars which do not have articulated enough means for multilayered linguistic representation (as this is possibly not even their

46For example amplius expressing the neuter nominative/accusative comparative form in the paradigm of AMPLUS ‘ample’ beside the masculine/feminine nominative amplior, accusative ampliorem, and being used in adverbial function as in ‘more amply’, while the expected regular amplioriter for adjectives of its inflection class is ungrammatical.
goal), and which for example do not distinguish between different grammaticalisation status of word classes, nor contemplate word class-changing inflection. We argue that this practice has undesirable consequences also on the representation of lexicon: in particular, its regular byproduct are unwarranted duos of words. This happens for example with the Latin form *primum*, interpreted either as an ADV ‘firstly’ with lemma *PRIMUM* when it has adverbial function, as in

(7)  [...] *cum primum pedes iuxta Sarni fluenta securus et incautus defigerem* [...]  
‘no sooner had I set foot by the streams of Arno, in all security and heedlessness’  
– Latin UDante Epi-35

or as an ADJ ‘first’ with lemma *PRIMUS* when e.g. it modifies a noun, as in

(8)  [...] *manifestum est quod celum primum magis recipit de luce primi* [...]  
‘it is manifest that the *first* heaven receives more of the light of the Prime’  
– Latin UDante Epi-346

But sometimes the annotation of the two words is identical in every other respect, like *weekly* in the two sentences

(9)  *I intend to provide *weekly* updates on the status of the above actions during April*  
– English ewt email-enronsent02_01-0060

(10) *the easiest thing would be a fish - but u need to clean the tank *weekly*, feed it the right amt, etc etc x*  
– English answers-20111108102204AAIvYN_ans-0006

tagged ADV and ADJ respectively, but with the same lemma *weekly* (and no morphological features besides a moot Degree=Pos when ADJ). These distinctions (cf. Table 4) appear questionable in a language’s vocabulary, as we are really talking about the same words, even if serving different syntactic functions, and they might be seen as a rather inconsistent annotation practice, against the *UD* desideratum of “rapid, consistent annotation” (de Marneffe et al., 2021, §5, point 3).

In sum, *pace* e.g. Pinkster (1972, §5), we argue that it is possible to claim that, with the goal of meaningful linguistic representation, transparently derived forms in adverbial function should be labelled just as all other forms belonging to their respective parts of speech, with all their pertinent morphological features, but possibly accompanied by a specific morphological feature marking the transposition. This feature would not, at any rate, be *Case* with a new value (here agreeing with Pinkster, 1972, §5), as it happens with the “adverbial case” in the *Index Thomisticus* (Busa, 1974–1980, p. xi): a transposition connects paradigms, instead of being constrained into one. Also, we argue that the counter-arguments against recognising word class-changing
inflection in the case of “adverbialisation” can easily be dealt with, if we consider the nature of the adverbial function of metapredication, as defined in Section 4.2. For instance, the fact that some properties of a part of speech seem to be blocked by an adverbial transposition, e.g. members of the NOUN class losing the ability to take adjectival modifiers (represented by relation amod), is directly explainable by observing that adnominal attribution cannot be applied to predications, but another adverb is needed instead, and that, at least in some languages like Latin, adverbial transposition seems to target single words and not phrases (unlike Case). The topic deserves a wider typological investigation, but, otherwise, it is actually not uncommon for adverbs to have arguments in the same way as their bases (cf. fn. 49 and Ricca, 2010, §7.2), as seen for example in

(11) Unfortunately for them, the general has decided to take seriously Colin Powell’s frequent boasts that it was on his nudging that the Indians made conciliatory gestures toward Islamabad.

– English eft
weblog.blogspot.com_dakbangla_20041028153019_ENG_20041028_153019-0025

Thus, we are probably in the presence of the peculiarity of a non-relational word class like NOUN as opposed to the others, rather than with a “strange” property of adverbial transposition. Members of the NOUN class are also remarkably less transposed into adverbs than ADJs (cf. §4.4.2 for Latin).

As already happens to verb forms by means of the morphological feature VerbForm (cf. Cecchini, 2021), in UD we propose to annotate marked deadjectival adverbs with a new similar feature, e.g. AdjForm=Adv. More in general, we can envision a generic pattern for such a feature, with name <POS>Form, depending on the part of speech of the base word, and values representing the external syntactic function of the transposition. For example, given the distinction for core adverbs put forward in Section 4.3, we could have DetForm=Part for determiners transposed into adverbs, as the previously mentioned Latin quantum ‘as many/much’. For participial adverbs like Latin diligenter ‘industriously’, as in

(12) Diligenter quippe notandum est quod dicit «nescit et nequit» [...] ‘It should be noted carefully that he says «neither knew nor could» [...]’

– Latin UDante Epi-365
glossable as dilig-ent-er ≡ diligo.IPPV.ACT.PTCP-ADV, from diligo ‘to esteem highly’, so lit. ‘highly esteemingly’, this treatment should be preferred to a more opaque

---

VerbForm=Conv. At the same time, this would need a layered treatment\(^48\) of this double transposition, for example as a recursive AdjForm[Part]=Adv|VerbForm=Part.\(^49\)

While members of ADJ and DET, followed by adjectival forms of VERBS, are the most common and direct sources of respectively autosemantic (ADV, traditionally manner adverbs, Schachter and Shopen, 2007, §1.4) and synsemantic (PART) adverbs, as they are already attributive and predicative in nature, one also encounters cases involving NOUNs like

- Latin viritim ‘man by man’, from NOUN vir ‘man’ (Ricca, 2010, §2.2.2), as in

  (13) Tribuniciae potestatis duodevicensimum, consul XII, trecentis et viginti milibus plebis urbanae sexagenos denarios viritim dedi

  ‘In the eighteenth year of my tribunician power and my twelfth consulship I gave 240 sesterces apiece to 320 000 members of the urban plebs.’

  – Latin Perseus phi1221.phi007.perseus-lat1.tb.xml@47

- Italian gattoni ‘on all fours’ from NOUN gatto ‘cat’ (thus literally ‘walking as a cat does’) with the derivational suffix -oni,

- or more productive German formations (Barz et al., 2009, §1161), such as BEISPIEL-SWEISE ‘for example’ and VORSICHTSHALBER ‘precautionally’, from NOUNs BEISPIEL ‘example’ and VORSICHT ‘caution’ respectively.

For these, a consistent annotation with NounForm=Adv might be felt as more controversial by some. Still, their annotation as transpositions, different from case forms, is justified when faced with the data at our disposal which do include case forms in adverbial function, and would be coherent with the proposal of this work; in fact, we suggest a unifying annotation along the lines of Transposed=ADV, which could take care of all cases of transpositions, e.g. Transposed=ADJ instead of VerbForm=Part, and so on. Similar detailed issues of annotation formalism, however, go outside the scope of this paper.

\(^{48}\)https://universaldependencies.org/u/overview/feat-layers.html

\(^{49}\)The effective difference between VerbForm=Conv (‘converb’, i.e. verbal form with the syntactic function of an adverb) and AdjForm=Adv plus VerbForm=Part (‘participle’, i.e. verbal form with the syntactic function of an adjective) deserves further investigations. We may suggest that the former is derived directly from a verb root and more easily maintains a verbal valence. However, we occasionally find cases like

(i) Premissis quoque rationibus consonanter […]

  lit. ‘thus agreeingly with the aforementioned reasonings […]’

  – Latin UDante Epi-351

where consona-nt-er = consona.IPFW.ACT.PTCP-ADV, from consona ‘to resound’, takes an argument (obl) in the dative case like the base verb. We notice that in Late Latin, with the disappearance of the so-called supine (effectively a converb, see Cecchini, 2021), this strategy seems the only one left to form adverbs from verb roots.
In the following Subsections, we present more in-depth discussions of adverbial transpositions and adverbs in general, to help complete the picture of how the word class of adverbs would look in UD annotation if implementing the framework of this work.

4.4.1. Two examples of transposition: Italian and Latin

We would like to illustrate through two interesting cases how a transpositional annotation shows its advantages on all annotation levels.

The first one is a case of attraction in Italian: in spontaneous production, when DET TROPPO ‘too many/much’ adverbially modifies another determiner, the former can pass from the “default” form troppo (singular masculine) to agreement with the latter, as in troppe poche volte instead of expected troppo poche volte ‘too few times’, where troppe, poche and volte are all feminine plural forms. If one does not want to annotate troppo/troppa always simply with lemma TROPPO, part of speech DET and relation advmod, accepting that it can function either as det or advmod in its bare form, one implicitly has to postulate the existence of four independent distinct members of ADV TROPPO, TROPPO, TROPPI and TROPPE. This is not an economic representation and does not seem to describe correctly what is happening at the morphosyntactic level, nor to allow for a simple, unified query of this phenomenon in the treebanks without the input of prior knowledge.

The second case consists in the interplay between so-called secondary predications and adverbs in a language like Latin (cf. Pinkster, 1990, §8, especially example (77), p. 155). In a sentence like

(14)  [Britanni] ex silvis rari propugnabant
‘The Britons made sorties from the woods in small groups’
– De Bello Gallico, Caesar

the form rari, from ADJ RARUS ‘thin, rare’, refers to Britanni ‘the Britons’, and as such agrees with it in case (nominative), gender (masculine) and number (plural). It is a secondary predicate or “optional depictive”, and according to UD’s guidelines it is annotated as advcl, as it forms an embedded (nominal or implicit) clause. This sentence is contrasted by the possible

50 For the following, cf. the discussion in https://accademidellacrusca.it/it/consulenza/troppopochi-nontroppipochi/1559 (only in Italian). Though based on a prescriptive point of view, it testifies to the spread of the phenomenon.

51 Cf. the guidelines for Latin (by the author) at https://universaldependencies.org/la/dep/advcl-pred.html, whence the example is taken.
(15) \[\text{[Britanni] ex silvis raro propugnabant}\]

‘The Britons \textit{seldom} made sorties from the woods’

where \textit{raro} ‘thinly, rarely, seldom’, singular masculine/neuter ablative of \textit{RARUS}, modifies the predicate \textit{propugnabant} ‘they were making sorties’ using a different strategy, i.e., a conventional case form, akin to what happens for so-called absolute constructions,\textsuperscript{52} since now it does not have to agree with any referent. Here we might also expect a more canonical \textit{rare} (adverbial form in \textit{-e}; see Ricca, 2010, §2.2.1). Annotating both instances \textit{rari} and \textit{raro} with the same lemma \textit{RARUS} and part of speech \textit{ADJ}, one with relation \textit{advcl:pred} and the other with \textit{advmod}, while keeping all usual morphological features, would be morphosyntactically motivated, look like the most economical representation and, moreover, allow for an easier, more “integrated” investigation of the distribution of an adjective like this one, without requiring particular prior knowledge from the querier.

4.4.2. What lies under \textit{advmod}: a focus on Latin

In this Subsection, we want to exemplify the mixed composition of elements with adverbial function and of the \textit{ADV} class, at the same time tackling the issues of the current annotation standards, by taking a closer look at what happens in \textit{ud} Latin treebanks. This brief overview is to be considered preliminary work for more thorough investigations about the treatment of Latin adverbs in linguistic annotation, it will be necessarily coarse-grained, and offers many pointers to the wider classification brought forth in Section 5.

Taking all Latin treebanks together,\textsuperscript{53} we retrieve 1360 unique lemmas for words annotated as belonging to the \textit{ADV} part of speech. After some minimal polishing of the data, which includes standard normalisation (lowercasing; neutralising the \textit{u/v} distinction), the expunction of non-lexical category labels (e.g. \textit{greek.expression, monetary}) and of compound forms (e.g. \textit{in+quantum, kal. maias}), and the selection of a canonical spelling when more than one is used (e.g. \textit{paulatim} and \textit{paullatim} retracted to \textit{paulatim} only), this number reduces down to 1285 unique \textit{ADV} lemmas. These still include some noise coming from spurious annotations, but can already deliver a meaningful picture.

We observe that as many as 918, that is 71.44\%, of these supposed \textit{ADV}s can be transparently traced back to autosemantic bases, as shown in Table 5. We notice that in our data some of these lemmas sometimes directly correspond to the base of the adver-

\textsuperscript{52}See \textit{ud} documentation for Latin (by the author) at https://universaldependencies.org/la/dep/advcl-abs.html and more generally (Pinkster, 1990, §7.4.7).

\textsuperscript{53}Refer to Paragraph \textbf{A note about data} in §1.
Table 5. Distribution by part of speech of autosemantic bases of ADV-tagged words in UD Latin treebanks.

<table>
<thead>
<tr>
<th>Part of speech of autosemantic base</th>
<th>Number of unique lemmas</th>
<th>Examples with bases</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADJ</td>
<td>608</td>
<td>PRINCIPALITER &lt; principalis ‘original’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ABSURDE &lt; absurdus ‘discordant’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PRIMO &lt; primus ‘first’</td>
</tr>
<tr>
<td>VERB</td>
<td>214</td>
<td>FESTINANTER &lt; festino ‘to hasten’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>IMMERTO &lt; mero ‘to deserve’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>DISTINCTIM &lt; distinguo ‘to separate’</td>
</tr>
<tr>
<td>NOUN</td>
<td>96</td>
<td>FUNDITUS &lt; fundus ‘bottom’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TABULATIM &lt; tabula ‘plot (of ground)’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NUMERO &lt; numerus ‘number’</td>
</tr>
</tbody>
</table>

Table 6. Presence of the different adverbial formations seen in Table 5 in UD Latin treebanks.

<table>
<thead>
<tr>
<th>Common ending</th>
<th>Number of unique lemmas</th>
<th>Distribution of parts of speech of base</th>
</tr>
</thead>
<tbody>
<tr>
<td>-e</td>
<td>315</td>
<td>236 ADJ, 79 VERB</td>
</tr>
<tr>
<td>-ter</td>
<td>234</td>
<td>173 ADJ, 61 VERB</td>
</tr>
<tr>
<td>-o</td>
<td>75</td>
<td>35 VERB, 30 ADJ, 10 NOUN</td>
</tr>
<tr>
<td>-im</td>
<td>36</td>
<td>22 VERB, 11 NOUN, 3 ADJ</td>
</tr>
<tr>
<td>-itus</td>
<td>8</td>
<td>4 NOUN, 4 ADJ</td>
</tr>
</tbody>
</table>

Adverbial form, e.g. AEQUALIS alongside of AQUALITER for the form aequaliter in Perseus, so we are actually dealing with an even lower number of effective lexemes. This shows oscillations in the manual annotation and/or conversion processes from other formats, which incidentally testifies how these adverbial forms are spontaneously felt part of an adjectival, verbal or nominal paradigm. In fact, all these formations look productive from the data at our disposal. In particular, we have rough distributions for these formations as shown in Table 6.\(^{54}\) No such formation is isolated, and each appears at every diachronic stage of Latin. Moreover, the data are not exhaustive, since we can find still more formations of the same kind in dictionaries, e.g. RADICITUS

\(^{54}\) We group them under common endings, refraining here from identifying the exact form of the suffixes operating on the bases.
from radix ‘root’ or frustatum from frustum ‘piece’ (see Lewis and Short 1879, s.vv.), showing their vitality. The distribution of the different formation strategies among parts of speech deserves further study, but it probably comes to no surprise that the most attested ones involve adjectival and verbal bases, and that the latter invariably pass through adjectival verb forms (“participles”, marked in UD as VerbForm=Part). Examples are festinanter from the imperfective festinans (so paraphrasable as ‘by hastening’), or distinctim from the perfective distinctum (so paraphrasable as ‘by having been separated (each from the other)’). At the same time, nominal bases are rarer and more limited in their strategies of adverbial transposition. This is in accordance with the fact that (adnominal) modifiers are more easily transposed into metapredications than non-relational word classes such as nouns, as discussed in Section 4.4.

It is also interesting to notice that a specialised lexical resource like Word Formation Latin (Litta and Passarotti, 2019) does not include the first two cases of Table 6, but instead includes words falling under the last three: the latter are not seen as part of paradigmatic inflections, while the former are. However, this choice seems dictated by an arbitrary cut based on frequency in the data, and is at odds, among others, with lexicographic conventions (see e.g. Gaffiot, 2016; Lewis and Short, 1879, s.v. absurde).

Some problems of contextual annotation (cf.§4.1) also arise from the data. One example, among the formations shown in Table 6, is the ending -o: this is actually the regular nominal ending for the singular ablative/dative masculine/neuter, and lemmas like those shown in Table 5 do not make exception. It is all the more striking, then, that Word Formation Latin singles them out as derived forms (through “conversion”), while their interpretation as ablatives of neuter forms is straightforward: primo ‘as first (thing)’, immerito ‘(this) not having been deserved’, numero ‘with number(s), i.e. precisely’. Then, there are also elements which should be better analysed as consisting of more than one word, such as saepenumero ‘oftentimes’ (< ADV saepe + NOUN numero) or superficietenus ‘up to the surface’ (< NOUN superficie + ADP tenus). This however opens up issues which are amply debated among linguists and in the UD community, and are too complex to be tackled here. What we rather want to point out is that the complex of similar “adverbialising” annotation choices brings about an unwarranted proliferation of adverbial lemmas and the fragmentation of lexemes. Forms like primo co-occur with primum and primitus, and distinctim with distincte. There might be reasons to distinguish some of them by means of extra morphological features (for example, the -im formations seem to express a distributorial meaning, possibly marked as NumType=Dist), but, from a paradigmatic point

55Some are also used for the corresponding synsemantic classes, such as qualiter < DET qualis ‘of what sort’, istim ‘thither’ < DET iste ‘this’.
56A so-called absolute construction; see fn. 52.
57As starting references, we can point to Haspelmath, 2017 and Lehmann, 2020, which incidentally considers many Latin examples.
of view, they all just correspond to the respective bases PRIMUS and DISTINGUO. A similar focused reduction (here, from 5 to 2 lemmas) would greatly decrease the actual number of 1285 unique lemmas, and this would improve both the representation of the lexicon and linguistic inquiries performed on the data, as discussed in Section 4.4.

Going beyond lemmas of ADVs with autosemantic bases, we can detect similar trends among the remaining 367 unique lemmas. Apart from compounds and obvious errors, many (238, the 18.52%) can actually be considered dubious annotations for forms better interpreted as pronominal ones, especially of the relative kind (e.g. ubi ‘where’; cf. §5.1.4), as conjunctions (e.g. SCONJ nisi ‘if not’, CCONJ et ‘and/too’; cf. §5.1.2 and §5.1.3), numerals (mostly multiplicative forms, e.g. septiens ‘seven times’ < septem ‘seven’), determiners (e.g. multo < multus ‘many, much’), adpositions (e.g. ab ‘from’), or interjections (uae ‘alas!’). Cursorily, it is interesting to observe that classes such as DET and PRON, synsemantic counterparts of ADJ and NOUN, seem to prefer inflectional strategies (especially accusative and ablative forms) instead of assuming specific adverbial forms (as represented by the -e and -ter endings in Table 5): this distinction gets lost with contextual annotation.

We are then left with a nucleus of 113 elements (8.79%). Among them, 51 represent the problematic class of “relators”, which in the treebanks typically alternate between the classes ADP and ADV, and also SCONJ, according to context, but which probably could be better treated as nominals (see §5.1.5). Examples are foris ‘outside’ and post ‘behind’. The remaining 72 (5.6%) elements, however, satisfy the condition of metapredicating (see §4.2) and of not being synchronically derivable from some base belonging to another part of speech. Of these, 26 show a clearly synsemantic nature: in our framework (see §4.3), they would be annotated as PART, as they already are in some treebanks: examples are the discoursive connective nam or the emphatic particle ecce. The core of autosemantic “true adverbs” consists then of around 46 elements: among them, mox ‘soon’, uix ‘with difficulty’, clam ‘secretly’, palam ‘publicly’, ferre ‘approximately’, semper ‘always’, comminus ‘at close quarters’. The unbalanced ratio between autosemantic and synsemantic elements is in line with the one seen for other part-of-speech dyads (cf. §3.2). The smaller size of the ADV class (46, or 97 including relators), as opposed for example to the much larger ADJ class (3025 unique lemmas across UD’s Latin treebanks) gives the impression of an “ancillary” word class: a function (metapredication) for which Latin, or any other language, has developed specific words, but which can usually be covered by other word classes, especially those of modifiers, deploying different strategies. This ancillarity might be one cause of the general underdefiniteness of the adverb class in traditional grammars of Latin, and in linguistics in general (cf. §1).

This in-depth look at Latin “adverbs” showcases phenomena which are present to a greater or lesser extent in every language of the world, each according to its mor-
phosyntactic characteristics. It also highlights widespread annotation approaches in UD and beyond. Our aim is again to demonstrate how the annotation framework presented in this work makes it possible to pursue typologically relevant inquiries, all the while rationalising the representation of lexicon and of the transpositional processes of the language. Many of the aspects touched upon in this Section will be discussed more in detail in Section 5.

4.5. Dependency relations of adverbs

Finally, it remains to be seen which dependency relations are best suited to adverbs and words acting with adverbial function. Given the split of core adverbs into two parts of speech along the autosemantic/synsemantic dichotomy (see §4.3), we might also envision to use a separate, specific dependency relation for members of PART with syntactic adverbial function, in the same spirit of distinguishing autosemantic amod and synsemantic det for autosemantic ADJ and synsemantic DET. On the other hand, this step might not be truly necessary, as the autosemantic/synsemantic dichotomy can be deduced just from the choice of part of speech (as it happens for the dyad NOUN/PRON).

If such a difference is to be implemented, we argue that the relation aux comes in handy here. The justification for this would be that synsemantic adverbs do not really specify attributes of a predication as defined in Section 4.2, as much as they convey more grammatical/pragmatical nuances, similarly as to how verbal AUXs “lend” Tense, Mood, Voice, etc. to a predication. Incidentally, we note that similar solutions are already adopted by other annotation formalisms, such as PROIEL for Latin and other ancient languages.\(^{58}\) Expanding the use of the aux dependency relation would also have the desirable consequence of making it typologically more interesting to track the morphosyntactic distribution of features inside a phrase, instead of keeping it redundantly locked to a specific part of speech (AUX).

5. A new taxonomy of adverbs, and beyond

This section deals with the internal restructuring of the word class of adverbs (ADV/PART), and especially with the repercussions that this has in relation to other parts of speech, given that the whole linguistic system is interconnected. The proposed new taxonomy is summarised in Tables 7 and 8. The former reinterprets UD’s part-of-speech system (Table 2) in view of the definitions and proposals of Section 4, and of the particular problems posed by some elements previously subsumed under the class ADV; part of it has to be considered tentative. The latter synthesises how

\(^{58}\)See the PROIEL annotation manual at \url{http://dev.syntacticus.org/annotation-guide/}, under the Section “Auxiliary words”.
non-core elements currently labelled as ADV should be redistributed among parts of speech, according to our framework.

We note that the proposals put forward in this Section can be seen as lying on a “gradient” from minor to possibly more radical (§5.1.5) interventions on the UD system of parts of speech, and/or readaptations of existing formalisms. While some of these changes might be easily implemented and, to some extent, could be considered to already exist in embryo among UD treebanks (such as ADJ taking the relation adv-mod in German, or today being annotated as NOUN in English), others might admittedly require more effort by the community, possibly involving a redefinition of annotation practices which are currently seen as established. As discussed in Section 3, the fundamental key point here is that if we do identify (morpho)syntactically different groups under the same label of ADV as it is currently intended, then it is not tenable to keep them together: in a way or the other, some action has to take place in order to be internally coherent with UD’s base framework. At the very least, a distinction between autosemantic and synsemantic elements inside the adverb class should be acknowledged. So, while Sections 4.3–4.5 form the most defined core of our proposal, and we think they should be seen as the cornerstone of a new, improved taxonomy of adverbs, this last Section might look sketchy in comparison. However, we note that it could hardly be any different, since, beyond a given point, many annotation choices will ultimately depend on language-specific considerations, so that here we can just hint to how the present framework can be followed down into all its ramifications, thereby leaving further definitions to future studies.

5.1. Better annotation strategies for current ADVs

This Section goes more into the detail of some particular groups of words that can be currently found in the ADV class, beyond the core adverbs as identified in Section 4. The treatment of some seems straightforward, but the more one reaches towards other “invariable” classes such as conjunctions and adpositions, the more notoriously difficult it seems to find a satisfying accommodation for some phenomena (cf. §5.1.4 and §5.1.5).
5.1.1. Not adverbs

In keeping with UD’s tripartite distinction for non-core dependents advmod/obl/advcl, according to a phrase’s head (de Marneffe et al., 2021, §2.3), some elements customarily labelled as ADV should simply be annotated by means of the oblique (obl) dependency relation. In other words, they do not belong to the ADV/PART class dyad, but rather to nominal (NOUN/PRON and possibly ADJ/DET/NUM), or even verbal (VERB and possibly ADJ/DET/NUM), classes.

Latin, having extensive morphology, is rich in such examples (some seen in §4.4.2). One of them is *forte* ‘by chance’: even if the base NOUN *fors* ‘chance’ is deemed defective,\(^\text{59}\) it still remains a NOUN (cf. Wackernagel, 2009, L. 149, p. 374, on *sponte* ‘of one’s free will’). Despite this, some Latin treebanks (PROIEL and UDante) uniformly tag this form as ADV.\(^\text{60}\) Analogously *primo* ‘at first’, from ADJ PRIMUS ‘first’, just a canonical use of the ablative case for quality or degree of difference (cf. Greenough et al., 2014, §398) in the sense of ‘as the first (thing)’ (also §4.4.2). We see similar occurrences in the German series *erstens* ‘firstly’, *zweitens* ‘secondly’, *drittens* ‘thirdly’, etc., all labelled as ADV in the German HDT treebank, but clearly related to *erst* ‘first’, *zweit* ‘second’, *dritte* ‘third’ as their genitive forms. In general, annotation should refrain from “adverbialising” nominal adjuncts (a case of contextual annotation; §4.1): they are

---

\(^{59}\)That is, lacking expected forms of its paradigm; in the case of *fors*, all but the singular nominative and ablative cases are attested.

\(^{60}\)Maybe influenced by lexical entries such as that in the Oxford Latin Dictionary (Glare, 2012, s. v.), which however admits the impossibility to distinguish an autonomous adverb.
already covered in ud by obliques. The difference between a case like *forte* and one like *raro* (see §4.4.1) is identifiable in that the former still regularly acts as an argument of the predicate, as we expect from members of the NOUN class, while the latter changes its attribution from a nominal (generic, implicit) referent to the predicate, but is not an argument itself. The interesting link (which gets lost with a redundant adverbial annotation) is though that both instances share the same strategy, here the use of the ablative case. The recognition that some of these forms can become “lexicalised” in time, or represent conventionalised, crystallised, possibly non-compositional expressions is part of a different, more lexicographical level of annotation, which does not necessarily coincide with the morphosyntactic one. Another cause for reflection is that similar “paradigm schisms” seem to be pursued only for non-core, oblique arguments, highlighting one of the dichotomies at the base of ud.

Finally, we note that this category might also include nominals which almost exclusively appear as oblique arguments: a typical case are temporal expressions like English *today*, Latin *cras* ‘tomorrow’. While the former has a transparent etymology, even if the second does not, they nonetheless both behave in the same way. This sub-category shows similarities to that of pro-forms, as discussed in Section 5.1.4, but it could be argued (actually contra 5.1.4) that these words display a semantic content which goes beyond mere deixis.

5.1.2. Focalisers

A particular subclass of synsemantic adverbs, this time part of the core ADV/PART dyad as defined in Section 4.3, are so-called focalisers or intensifiers, currently tied to the subtyped relation advmod:emph (§3.1.1), which in the schema presented here could possibly become aux:emph (cf. §4.5). They are peculiar in having a wider distribution than autosemantic adverbs, often including nominals in their scope. From ud’s guidelines:

“While other adverbial modifiers usually modify verbs, adjectives or adverbs, these emphasizers often modify noun phrases, including prepositional phrases.”

From Table 3 and data in §3.2 we can see them appear with much higher frequencies than other ADV-labelled elements, a Zipfian distribution typical of grammatical words: so in German *auch* ‘also’; in Latin *etiam* ‘also’ (39.7% of its occurrences modifying a non-copular nominal in IT-TB, the most among significantly frequent adverbs together with *solum* ‘only’), *sic* ‘so’, *taliter* ‘in such a way’, *tantum* ‘so much’. The last three

---

61 Cf. Gamba and Zeman, 2023, p. 10 for an overview of other similar cases in need of harmonisation in Latin treebanks. This has started taking place with release v2.13.

62 But see Savary et al. (2015) for a project dealing with a similar topic, possibly to be integrated into ud in the future (Savary et al., 2023).
terms all are demonstrative elements (PronType=Dem), and the last two are in fact
derived from determiners DET **talis** ‘of that kind’ and DET **tantus** ‘that many/much’,
making their synsemantic nature explicit. We also note that, prior to UD v2.12, in
Latin Perseus **etiam** was grouped with CCONJ.s. In fact, co-ordinating conjunctions,
also unrestricted in distribution, do regularly function as focalisers, like Latin **et** ‘and’
(7.6% of its occurrences in the IT-TB) or **neque** ‘and not’. This topic deserves further
studies, but here we can just suggest that the identification of focalisers as synsemantic
adverbs (PART, aux:emph) can be justified in that, even though they do not directly
metapredicate (§4.2), the focus they supply to a noun phrase has a meaning only in
relation to the predicate itself (differently than, say, an ADJ): e.g. in

(16)    _et vix Ytali infelices lacrimis metiuntur_
‘even the Italians in their woe can scarce measure with their tears’
    – Latin UDante Epi. - 79

the _et_ emphasises how _Ytali_ as nominal subjects (nsubj) perform the action. Focalisers
thus metapredicate indirectly through the argument of a predicate, and this way we
explain also the rarer, occasional adverbial modification of nouns by autosemantic or
pronominal adverbs (cf. Pinkster, 1972, §4.3.2f. for Latin). Negations as members
of PART, as mandated by the UD guidelines, also enter in this picture, in that they are
a form of focus by contrast of a given constituent.

5.1.3. Discourse and connecting elements

Some adverbs seem to metapredicate an entire sentence, i.e. they are somehow
“external” to the rest of a clause (“modal” or “evaluative” adverbs; for Latin see Ricca,
2010, §4; Pinkster, 1972, §6.2.1.1, speaks of “periphery”). Other elements are seen to
connect clauses (Kortmann, 1997; Rosén, 2010), but they move more freely than con-
junctions and independently of them, even if the two classes can be confused, or a mixed,
distinct category considered, like _Konjunktionaladverbs_ in German for words
such as **trotzdem** ‘despite’ (Barz et al., 2009, §864).

The common theme to all these elements seems to be textual coherence, expressed
outside the argumental structure of a clause, and so the relation discourse might
possibly suit them better than advmod (or aux). The exact nature of these connect-
tives would be given by their adverbal subclass: ADV or another autosemantic basis

---

63 As in Lewis and Short (1879), but not in Glare (2012). For the harmonisation in UD v2.12 see Gamba
and Zeman (2023).

64 We note that UD is not directly concerned, with regard to annotation, with possibly systematic semantic
variations, as seen for example for exactly veering towards an intensifier (§3).

65 So, for Latin, in Lewis and Short (1879) the particle _nam_ is listed as a conjunction, while other sources
such as Glare (2012) retain it a “particle” (an even less defined term than “adverb” in this context).
for elements effectively steering the meaning of a sentence, e.g. Greek fysiká ‘naturally; of course’, from ADJ fysikós ‘natural’, when not taken literally, as opposed to the semantically completely bleached and non-derived Latin quidem (Kroon, 2005; already annotated as PART & discourse in some treebanks), which shows a clitic-like, Wackernagelian behaviour in fixed second position, as many other similar elements (Wackernagel, 1892; Spencer and Luís, 2012, §3.2.1). If these elements are effectively distinguishable from conjunctions, or could possibly be annotated as such (by means of cc) while retaining an adverbial part of speech, remains an open issue.

5.1.4. “Pronominal adverbs”

uo’s guidelines for ADV make a peculiar distinction for “pronominal adverbs”, and a special position is given them also in Zeman (2018, §3.4, Table 3.15). Since they are defined as “refer[ring] to circumstances in context, rather than naming them directly”, which reads as deixis (see Anderson et al., 2003, s.v.) (i.e. a function typically associated to the class PRON and/or DET), there seem to be very good reasons for not annotating them as ADV. Indeed, uo’s instructions are to analyse them as pronouns/determiners under all respects (e.g. they take the feature PronType) apart from part of speech and, consequently, dependency relation. There are actually many factors in favour of an analysis of such elements simply as pronouns (PRON):

1. a parallelism with the oblique arguments they substitute. This needs the acknowledgment that there can be oblique pronouns (as there apparently are oblique nouns, cf. §5.1.1), which give the impression of being “invariable” (as recounted by many traditional grammars) only if taken in isolation, but are possibly clearly part of paradigms, such as Latin hic ‘here’ descending from DET HIC ‘this (one)’ (declined for case, gender and number as hic, haec, hoc, huius,…);

2. as a consequence, this would eliminate purely morphology-based asymmetries such as Latin quis ‘who’ (inflectable) tagged as PRON, but QUANDO ‘when’ (“uninflectable”, but clearly an old ablative form tied to qui ‘what, which’, cf. de Vaan, 2008, s.v.) as ADV (both with PronType=Int or PronType=Rel), which becomes striking when one considers the fact that they both substitute for nominal phrases in relative clauses.

3. these elements appear in exactly the same contexts as adjuncts, sometimes even accompanied by adpositions (case) as any member of NOUN;

---

66 It seemed to be the case in Bulgarian for elements such as KĂDE ‘where’, KOGA ‘when’, KAK ‘what’, as reported by Zeman (2018, p.40), but they currently (v2.13) all appear tagged with ADV in UD.

67 Consider for one example 37 in the discussion about “adverbial relativizers” in the context of enhanced dependencies at https://universaldependencies.org/u/overview/enhanced-syntax.html#relative-clauses, given the same context, where and the episode act and should be annotated the same way, i.e. as nominals with relation obl. This would be valid for all so-called “relative adverbs”.
4. deixis, which is compatible with copular constructions as in

(17)  
\[ I \textit{was there} \]
– English ewt reviews-374000-0004

counter to the behaviour of core adverbs.

5. the possibility to have adnominal dependents like relative clauses (\texttt{acl:relcl}), as in Greek

(18)  
\[ \textit{ekei ópou den eínai dynatí} […] \]
\textit{‘there where […] is not possible’}
– Greek gdt gdt-20020206-ep-sessions_174-14

6. their strong if not exclusive preference for co-ordination among themselves (Pinkster, 1972, §7.3.1).

Under the same light, we can subsume under \texttt{PRON} also other elements traditionally considered adverbs like Latin reciprocal \textit{invicem} ‘one another’, or English \textit{yes} and \textit{no}, substituting for whole phrases and clauses. In general, this perspective entails considering \texttt{PRON}’s class not limited to personal pronouns and/or inflectable elements, but including pro-forms in general (see Anderson et al., 2003, s.vv. Pro-Adverb and Pro-Form), and possibly also taking into account other prepositional-like elements which have always been problematic for a dependency analysis (see next Section). What is common to all of them is that they realise arguments inside a predication instead of metapredicating it.

5.1.5. Relators and adpositions

In many (at least Indo-European) languages, for some elements traditionally considered adverbs for their “signifying time or place”, an apparent oscillation between fully autosemantic and adpositional (and thus synsemantic) behaviour is observed (for Latin, see Ricca, 2010, §3.2.1, Pinkster, 1972, §9, and Pinkster, 1990, §5). An example is Latin \texttt{procul} ‘at a distance’, occurring both totally independently from nominal elements, as in

(19)  
\[ \textit{cum […] id frater […] procul animadvertisset} \]
\textit{‘when the brother saw this from afar’}
– Latin proiel 53360

or with its own arguments, sometimes introduced by “proper” prepositions, as in

---

68For example, in English ewt there are no elements besides pronominal or relator (see §5.1.5) ones like \textit{here}, \textit{there}, \textit{how}, \textit{out}, \textit{where} etc. in a copular construction out of 287 occurrences of an \texttt{ADV} with a cop dependent.
(20) *non est autem procul a uero quod [*]

‘It is not, of course, far from true to say that [*]’

– Latin

Similar behaviour is shown e.g. by *above* in English or *mésa* ‘inside’ in Greek, as opposed to non-independently appearing words such as English *to*.

While agreeing with Gerdes and Kahane (2016, §4.1) that a “fracture” in the annotation of these two allegedly different construction types is undesirable, and not completely justified, it seems possible to argue that these kinds of elements, sometimes called relators, are still subordinated to their more autosemantic accompanying terms (if any). On the contrary, it does not appear to be useful to treat them as heads as proposed in Gerdes et al. (2018); Osborne and Gerdes (2019), or also Vincent and Börjars (2020), or at least this appears to run counter to the primacy of content words so fundamental in UD (de Marneffe et al., 2021, §2.2.1). Nonetheless, it is not easy to find a solution to this *vexata quaestio*, but in the framework of this paper we can envision two, possibly controversial, solutions, besides living them in the class *ADV* as is currently the case.

One could be to acknowledge these elements, too, as pro-forms specialised in particular spatio-temporal relations. Thus, they would be treated not too dissimilarly from deictic elements like *there* (cf. 5.1.4), having in common a context-dependent, so not totally autosemantic, meaning, and would side with *PRON*. Under this light, this category could be considered (*contra* §5.1.1) to contain also terms like Latin *cras* ‘tomorrow’ or German *gestern* ‘yesterday’: these point neither to a specific day like *Monday*, nor to a generic idea like *day*, but acquire meaning only in relation to the time of utterance. In contrast to *procul* or *above*, though, these words do not seem to take their own arguments. In any case, they differ from a core adverb (*ADV*) like Latin *saepe* ‘frequently’, which has a “fixed”, context-independent meaning (cf. Ricca, 2010, §3.1.2). Therefore, in a phrase like the previous example *procul a uero*, a *PRON* *procul* can be assumed to depend on *uero* (from ADJ *verus* ‘true’), possibly with a case relation, not so distant from pronouns acting as determiners (*det*); at the same time, it would be perfectly fit for a pronoun to appear independently.

A second, more radical solution (that we actually favour) could start from the identification of a spectrum of grammaticalisation between “adverbial” and “true” prepositions, and could thus advocate the introduction of a new part of speech, say *REL*, forming a dyad of relators with *ADP*. This class could even be considered to extend to other connectives, i.e. to conjunctions *CCONJ* and *SCONJ*. A kind of unitary analysis

---

69 Note that similar terms are often derived from semantically bleached noun phrases, such as Latin *home* ‘today’ from *hoc die* ‘(on) this day’ (de Vaan, 2008, s.v.), exactly equivalent to the corresponding English term.
is already present in UD with the case/mark alternation for adpositions (ADP) introducing respectively noun phrases and clauses.

Admittedly, these phenomena are very complex and a complete treatment of them lies outside of the scope of the present paper, so that the previous suggestions have to be left as tentative and hopefully as a stimulus for further research.

6. Conclusions and perspectives

The topic touched upon in this paper is of vast scope: the definition of adverbs in general, and the cases presented in §5 are each dealt with by a rich literature, branching out and intertwining with many other subjects in linguistics. This means that, beyond setting the main points for an improved adverbial annotation schema in UD and beyond (§3 and §4), the discussion in a work like the present one inevitably must be concise when it comes to specific cases (as in §5). This notwithstanding, the author’s hope is that the need for “tidying up” a neglected, traditionally poorly defined (§1) part of speech has been convincingly motivated and brought (maybe not for the first time) to the attention of the community. Further, it can be asserted that theoretical and practical problems raised by a more systematic treatment of UD’s ADV class can have a positive impact in contributing to an ever more consistent and universal annotation style, also beyond adverbs themselves.

Acknowledgements

I wish to thank the anonymous reviewers who helped to shape a chaotic heap of ideas into something which, I hope, is now understandable. Also, I thank the community of annotators for the invaluable contributes that the confrontation of ideas brings about every day.
Bibliography


Hajič, Jan, Eva Hajičová, Marie Mikulová, and Jiří Mírovský. Prague Dependency Treebank. In Ide, Nancy and James Pustejovsky, editors, Handbook of Linguistic Annotation,


Zeman, Daniel, Martin Popel, Milan Straka, Jan Hajic, Joakim Nivre, Filip Ginter, Juhani Lusatolahti, Samp Pyyealo, Slav Petrov, Martin Potthast, Francis Tyers, Elena Badmaeva, Memduh Gokirmak, Anna Nedoluzhko, Silvie Cinkova, Jan Hajic jr., Jaroslava Hvlovacova, Václava Kettnerová, Zdenka Uresova, Jenna Kanerva, Stina Ojala, Anna Missilä, Christopher D. Manning, Sebastian Schuster, Siva Reddy, Dima Taji, Nizar Habash, Herman Leung, Marie-Catherine de Marneffe, Manuela Sanguinetti, Maria Simi, Hiroshi Kanayama, Valeria dePaiva, Kira Dropogova, Héctor Martínez Alonso, Çağrı Çöltekin, Umutsulubacak, Hans Uszkoreit, Vivien Macketanz, Aljoscha Burchardt, Kim Harris, Katrin Marheinecke, Georg Rehm, Tolga Kayadelen, Mohammed Attia, Ali Elkahky, Zhuoran Yu, Emily Pitler, Saran Lertpradit, Michael Mandl, Jesse Kirchner, Hector Fernandez Alcalde, Jana Strnadová, Anna Melnikova, Katia Aplonova, Angelia Aquino, Carolina Aragon, Gryd Aranes, Maria Jesus Aranzabe, Bilge Nas Arican, Pörunn Arnardottir, Gashaw Arutie, Jessica Naraiswari Arwidarasti, Masayuki Asahara, Katla Ásgeirdstottir,

Address for correspondence:
Flavio Massimiliano Cecchini
flaviomassimiliano.cecchini@kuleuven.be
Erasmushuis
Blijde-Inkomststraat 21
3000 Leuven
Belgium
INSTRUCTIONS FOR AUTHORS

Manuscripts are welcome provided that they have not yet been published elsewhere and that they bring some interesting and new insights contributing to the broad field of computational linguistics in any of its aspects, or of linguistic theory. The submitted articles may be:

- long articles with completed, wide-impact research results both theoretical and practical, and/or new formalisms for linguistic analysis and their implementation and application on linguistic data sets, or
- short or long articles that are abstracts or extracts of Master’s and PhD thesis, with the most interesting and/or promising results described. Also
- short or long articles looking forward that base their views on proper and deep analysis of the current situation in various subjects within the field are invited, as well as
- short articles about current advanced research of both theoretical and applied nature, with very specific (and perhaps narrow, but well-defined) target goal in all areas of language and speech processing, to give the opportunity to junior researchers to publish as soon as possible;
- short articles that contain contraversing, polemic or otherwise unusual views, supported by some experimental evidence but not necessarily evaluated in the usual sense are also welcome.

The recommended length of long article is 12–30 pages and of short paper is 6–15 pages.

The copyright of papers accepted for publication remains with the author. The editors reserve the right to make editorial revisions but these revisions and changes have to be approved by the author(s). Book reviews and short book notices are also appreciated.

The manuscripts are reviewed by 2 independent reviewers, at least one of them being a member of the international Editorial Board.

Authors receive a printed copy of the relevant issue of the PBML together with the original pdf files.

The guidelines for the technical shape of the contributions are found on the website [https://ufal.mff.cuni.cz/pbml](https://ufal.mff.cuni.cz/pbml). If there are any technical problems, please contact the editorial staff at pbml@ufal.mff.cuni.cz.