Towards automatically extracting morphosyntactical error patterns from L1-L2 parallel dependency treebanks

Arianna Masciolini and **Elena Volodina** and **Dana Dannélls**
Språkbanken Text
Department of Swedish, Multilingualism, Language Technology
University of Gothenburg
firstname.lastname@gu.se

Abstract

L1-L2 parallel dependency treebanks are UD-annotated corpora of learner sentences paired with correction hypotheses. Automatic morphosyntactical annotation has the potential to remove the need for explicit manual error tagging and improve interoperability, but makes it more challenging to locate grammatical errors in the resulting datasets. We therefore propose a novel method for automatically extracting morphosyntactical error patterns and perform a preliminary bilingual evaluation of its first implementation through a similar example retrieval task. The resulting pipeline is also available as a prototype CALL application.

1 Introduction

L1-L2 parallel dependency treebanks are corpora where sentences produced by learners of a second language (L2), paired with native-like (L1) correction hypotheses, are annotated following the Universal Dependencies (UD) standard (Nivre et al., 2020). This data format, proposed by Lee et al. (2017), has interoperability as its main goal: UD provides a uniform annotation layer across different languages and its fine-grained morphosyntactical analysis is meant to make explicit error tagging unnecessary, preventing the incompatibilities that arise from the use of project-specific taxonomies. In addition, the availability of increasingly reliable dependency parsers can significantly speed up, if not completely automate, the annotation process.

Putting L1-L2 treebanks into use, however, requires effective ways to extract information from them. Errors, explicitly marked in most learner corpora, are for instance not straightforward to identify in such datasets. In this paper, we report on ongoing work on this problem, focusing on morphosyntax. In particular, we propose a novel approach to locate error-correction pairs and convert them into machine-readable error patterns, which can serve as a starting point for a variety of tasks, including explainable automatic error classification and controlled feedback comment generation.

We put a first implementation of this method to the test through an example retrieval task where patterns extracted from a set of example sentence-correction pairs are used to find similar errors in an L1-L2 treebank. An interactive version of the resulting system is also made available as a prototype Computer-Assisted Language Learning (CALL) application, similar to Arai et al. (2019)'s corpus search tool for L2 Japanese learners.[1]

2 Related work

Standardizing and automating the annotation of learner corpora is desirable for a variety of purposes. Notable in this sense is ERRANT (Bryant et al., 2017), an automatic ERRor ANnotation Toolkit for learner English whose principal aim is allowing finer-grained evaluation of Grammatical Error Correction (GEC) and Detection (GED) systems. ERRANT extracts edit operations from learner sentence-correction pairs. Each edit is later labelled following an error taxonomy relying solely on dataset-agnostic information such as the POS (Part Of Speech) tag of the tokens involved.

With L1-L2 parallel UD treebanks, there is no explicit error annotation step: the idea is that morphosyntactical annotation should suffice, as error can be described by means of tree patterns pairs, comparing the original learner attempt with its target L1 counterpart (Lee et al., 2017). When it comes to retrieving instances of specific patterns of error, a query engine was developed by Masciolini (2023). Choshen et al. (2020), on the other hand, used UD-annotated parallel data to automatically derive SERCL, a new taxonomy of Syntactic ERrors for automatic CLassification, later combined with ERRANT's under the name of SERRANT (Choshen et al., 2021). SERCL error types

[1] Our software is available for download at github.com/harisont/L2-UD (accessed 31.05.2023).

Figure 1: A correct-incorrect UD sentence pair both in English and Swedish, with discrepancies highlighted in bold.

are obtained by concatenating the morphosyntactical features of the head of a problematic text segment before and after correction. The results are labels such as ADJ→ADV (adjective replaced by adverb), applicable for instance to the example in Figure 1. Choshen et al. (2020)'s system, as well as the query tool, has been tested both on manually annotated treebanks and on automatically parsed sentences, with results suggesting the standard parsers' relative robustness to learner errors.

Querying parallel UD treebanks and using them to automatically derive data-driven error taxonomies are two tasks closely related but not identical to what we attempt in this paper. As opposed to searching for specific error types, we try to detect all errors appearing in an L1-L2 treebank, and rather than classifying them according to a flat labelling scheme we aim at obtaining fine-grained descriptions of each, in the form of patterns meant for further processing.

3 Methodology

We see error pattern extraction as a two-stage process. Given a learner sentence and the corresponding correction, the first step, discussed in Section 3.1, is locating its problematic portions to extract error-correction pairs. As per Section 3.2, the latter are then converted into machine-readable patterns.

3.1 Locating error-correction pairs

A simple way to locate errors in a pair of sentences is to phrase- and/or word-align them and consider as erroneous all correspondences presenting any discrepancies between their L1 and L2 components. If the goal is to only select errors belonging to a specific macro-category, the task of deciding whether a discrepant alignment is relevant or not becomes less straightforward. In this case, we are mostly interested in morphosyntax, for which UD annotation is particularly informative. At this stage, however, we assume our data to only contain this type of errors and focus on alignment alone.

That of alignment is a problem common to all the works mentioned in Section 2. To extract edits, ERRANT uses a linguistically-enhanced L1-L2 algorithm (Felice et al., 2016). While reportedly achieving state-of-the-art results, its implementation is English-specific. Choshen et al. (2020), on the other hand, work in a bilingual setting. The paper leaves the details of the alignment step unspecified, but from a superficial inspection of the source code it appears that the same method, along with an *ad-hoc* adaptation to Russian, is used.

Since our aim is to work cross-lingually, we adopt the same approach as Masciolini (2023), consisting in extracting correspondences between UD subtrees using the CONCEPT-ALIGNMENT package (Masciolini and Ranta, 2021). Originally developed for the syntax-based extraction of translation equivalents from multilingual parallel UD treebanks, the library is completely language-agnostic at its core, and its alignment rules can be easily customized to better suit the L1-L2 domain.

Furthermore, extracting subtrees rather than text spans ensures some degree of flexibility in determining how much context to extract for a given error. Depending on the use case, error-correction pairs can consist either of just the tokens involved in the corresponding edit operation, similarly to what is done in SERRANT, or of larger segments, useful to understand why the edit is required. In Figure 1, for instance, both the adverb *slowly* and the adjective *slow* (resp. *långsamt* and *långsam*) are acceptable forms, if taken in isolation: adjectives are only marked as incorrect because they modify a verb. For each detected error, our extraction module produces patterns of various sizes. From the perspective of example retrieval, in fact, smaller patterns are more likely to generate hits, but larger ones result in better matches.

3.2 From CoNLL-U trees to error patterns

Alignments, and therefore errors, are internally represented as pairs of *rose trees*, tree structures with a variable, unbounded number of children per node. While this representation can be easily converted back into CoNNL-U format, which is itself machine-readable, complete UD sentences are too information-rich for most practical purposes and not as easy to manipulate as a recursive data struc-

ture. We therefore describe errors using a UD query language. Among several existing options, we selected the pattern matching language available as part of GF-UD (Kolachina and Ranta, 2016; Ranta and Kolachina, 2017), the easiest to integrate with the rest of the codebase.

UD patterns GF-UD essentially provides three types of patterns:[2]

- *single-token patterns*, such as `POS "ADJ"`, matching subtree roots. With a similar syntax, it is possible to pattern match based on the token's XPOS, DEPREL, FEATS, FORM or LEMMA, each corresponding a CoNNL-U field[3];
- *tree patterns* in the form `TREE p [ps]`, where `p` is a pattern to be matched by the root of a subtree and `[ps]` a list of patterns denoting its dependents. `TREE (POS "NOUN") [DEPREL "amod"]`, for instance, matches nouns modified by an adjective;
- *sequence patterns* like `SEQUENCE [DEPREL "amod", POS "NOUN"]`, matching nouns preceded by an adjectival modifier.

In addition, the language allows combining patterns with the logical operators AND, OR and NOT and provides a TRUE pattern matching any subtree.

Following Masciolini (2023), we use pairs of these UD patterns to describe the discrepancies between L1 and L2 trees. As a consequence, a way to describe the error in Figure 1 on the basis of POS tags is the following:[4]

```
⟨TREE_ (POS "VERB") [POS "ADV"],
 TREE_ (POS "VERB") [POS "ADJ"]⟩
```

Here, the first pattern denotes the correct form and the second the erroneous learner attempt. This can be written even more concisely as

```
TREE_ (POS "VERB") [POS {"ADV"→"ADJ"}]
```

This means that, to modify a verb, the learner used an adjective rather than an adverb. If we focus on the edit operation only, we obtain the pattern

```
POS {"ADV"→"ADJ"}
```

equivalent to SERCL/SERRANT's ADJ→ADV.

[2]For the full specification of the GF-UD pattern syntax, see github.com/GrammaticalFramework/gf-ud/blob/master/doc/patterns.md (accessed 19.04.2023).

[3]For more information about the UD standard, see universaldependencies.org (accessed 31.05.2023).

[4]Underscored TREE_ patterns match even trees having dependents other than those explicitly listed, like Figure 1's.

Converting alignments to tree pattern pairs, which have the same recursive structure, is extremely simple. The same can be said of sequence patterns, since GF-UD also provides a list-like data type to represent UD sentences and functions to convert between the latter and rose trees. The most straightforward approach, however, yields "full" UD patterns that are excessively specific. For this reason, we develop various simplification strategies producing more general, yet informative patterns.

Simplification strategies A first, simple strategy, is to **filter patterns by CoNNL-U field**. This was already exemplified above when only considering Universal POS tags. A less strict options is to take into account all morphosyntactically relevant fields (FEATS, DEPREL, POS and possibly XPOS). A way to achieve further simplification is to **remove fields whose values are identical in both components of the patterns**. Another approach is to recursively compare the L1 and L2 sides of an error pattern and **eliminate identical subpatterns**. In addition, it is possible to **simplify single (monolingual) patterns** in various ways, for instance by transforming sequence patterns of length 1 and tree patterns with empty dependent lists into single-token patterns. Appendix A demonstrates the application of these strategies to the example in Figure 1. With example retrieval in mind, we apply all strategies, in sequence, to each extracted pattern, without discarding the intermediate results. This maximizes the chance of finding relevant examples while laying the foundation for ranking the results.

4 Preliminary evaluation

We carry out a first evaluation of our method through an example retrieval task. In particular, we try to find occurrences of errors similar to those extracted from a given sentence-correction pair in an L1-L2 treebank. Implementation-wise, this is done by combining our error extraction module with Masciolini (2023)'s query engine: run on an input pair, the extraction procedure returns one or more patterns, in turn used to query the treebank.

We make an interactive version of such error retrieval pipeline also available as a prototype CALL application, analogous to the incorrect example retrieval tool presented in Arai et al. (2019). In this case, input sentences are entered as text and parsed on the fly using UDPipe's REST API.[5]

[5]lindat.mff.cuni.cz/services/udpipe/api-reference.php (accessed 31.05.2023).

4.1 Data

While the final iteration of our extraction method will be meant for authentic learner data, we carry out this first evaluation on two datasets for linguistic acceptability judgments composed of minimal correct-incorrect sentence pairs isolating specific linguistic phenomena, i.e. where the incorrect element contains a single grammatical error. In this way, we postpone dealing with the complexities that can arise from the simultaneous presence of several errors involving the same tokens. We simplify the task further by filtering out sentences containing errors beyond mere morphosyntax, such as incorrect lexical choices and spelling mistakes, for which automatic UD annotation is less informative and potentially misleading.

BLiMP The Benchmark of LInguistic Minimal Pairs (BLiMP) (Warstadt et al., 2020), developed for evaluating the linguistic knowledge of language models, is a dataset consisting of 67 subsets, each containing 1 000 correct-incorrect sentence pairs exemplifying a specific error type or *paradigm*. Examples are artificially generated based on linguist-crafted templates and subsets are organized in 12 groups on the basis of the linguistic phenomenon they describe. Based on their metadata, we select lexically identical pairs marked as belonging to the fields of morphology or syntax and parse them with UDPipe 2 (Straka, 2018)'s default English model. The result is a parallel treebank of 14 996 sentences, 100 of which we set aside as inputs for the example retrieval pipeline. Specifically, we extract patterns from this 100-sentence subset and match them against the remaining 14 896 pairs to retrieve similar correct-incorrect examples.[6]

DaLAJ The DAtaset for Linguistic Acceptability Judgments (DaLAJ) is, in turn, composed of L2 Swedish sentence-correction minimal pairs derived from the error-annotated SweLL SWEdish Language Learner corpus (Volodina et al., 2019) and therefore arguably closer to the data our system is being built for.[7] SweLL uses a two-level error taxonomy: labels, such as M-Adj/adv, are composed of a capital letter, indicating the error's macro-category (in this case, Morphology), followed by an abbreviation specifying the affected POS and/or morphological features. The M-Adj/adv label, for instance, refers to Adjective forms corrected with the corresponding adverb, such as *långsamt* → *långsam** in the example displayed in Figure 1. We select the 1 198 error-correction pairs belonging to the M and S macro-categories and process them analogously to BLiMP data, the only difference being the usage of a Swedish model.[8]

4.2 Results

Ideally, quantitatively evaluating the performance of our system on the example retrieval task defined above would involve computing the precision and recall of each query performed with the extracted patterns. In practice, however, this is unfeasible in our current setup, as it would require manually inspecting all matches. While an identity of error labels between the input pair and a match is generally a good indication of a true positive, in fact, it is not at all always the case that different labels correspond to a false positive: the same error can sometimes be interpreted, and therefore labelled, differently. The Swedish word *långsamt*, for instance, is both an adverb ("slowly") and the singular neuter form of the adjective *långsam* ("slow"), meaning that a phrase like *ett {långsamt → långsam*} tempo* ("a slow tempo", where *{långsamt → långsam*}* modifies the neuter noun *tempo*) could, following the SweLL annotation guidelines (Rudebeck and Sundberg, 2021), be annotated both as M-Adj/adv and M-Gend. For similar reasons, counting actual false negatives is also challenging.

Instead, for each dataset, we compute the retrieval rate R, i.e. the percentage of sentences for which the system was able to return one or more matches, regardless of their correctness, and compare it with the *successful retrieval rate* R_+, where only sentences with at least one relevant match was found. Since we use search results as a proxy of the usefulness of the extracted patterns rather than to assess the performance of the query engine, we deem this to be sufficient for a first evaluation. Results are summarized in the table below.

	BLiMP	**DaLAJ**
R	82%	69%
R_+	82%	63%

[6] The BLiMP splits used in this paper, as well as the preprocessing scripts, are available at github.com/harisont/L1-L2-BLiMP/tree/bea (accessed 31.05.2023).

[7] An early version of DaLAJ, covering only lexical errors, is presented in Volodina et al. (2021).

[8] The DaLAJ splits used in this paper, as well as the preprocessing scripts, are available at github.com/harisont/L1-L2-DaLAJ/tree/bea (accessed 31.05.2023).

Figures for BLIMP, whose data is controlled and finely categorized by paradigm, were obtained fully automatically by checking whether one or more of the retrieved examples belonged to the same subset. DALAJ matches, on the other hand, still required manual inspection due to the dataset's coarser-grained labelling scheme and the scarcer predictability of the sentences. More specifically, we checked the search results of each query looking for relevant matches, defined, for the sake of this evaluation, as examples presenting an error similar to that of the input pair, regardless of the degree of specificity and granularity of the extracted pattern(s). Given the input *de blev {utsatta → utsattad*} på två olika sätt* ("they were exposed in two different ways", where the adjective *utsattad**, "exposed", is incorrectly inflected for number), for instance, this implied considering the sentences *{promenader → promenad*} är bra för människors hälsa* ("{walks → walk*} are good for people's health", where the number inflection error involves the noun) and *vi är {glada → glad*} varje dag* ("we are happy every day", where the incorrectly inflected word is again an adjective, *glad*) even though only the latter involves the same POS[9]. While results are encouraging for both datasets, we observe a marked difference between the two in terms of retrieval rate. Several different factors might contribute to this: the difference in size between the two corpora, the fact that all pairs we selected from BLIMP, but not from DALAJ, are lexically identical and some intrinsic characteristics of the BLIMP dataset, such as the template-based method used to generate its sentences.

In cases where no or exclusively incorrect matches are found, failures may also be caused by parse errors, issues related with the query engine or, especially when it comes to the smaller Swedish treebank, merely by a lack of similar examples in the corpus. In such instances, we investigate further by inspecting the UD trees and extracted patterns. When it comes to BLIMP data, pairs with no matches belong in all but one case to the island effects group, comprising word order errors related to wh-words, such as *Whose {hat should Tonya wear → should Tonya wear hat*}?* Unsurprisingly, errors of this kind pose a challenge for the parser and therefore often incorrectly aligned.

Word order errors are problematic in Swedish too, but even other syntactical errors, most notably S-Clause (change of basic clause structure), S-MSubj (missing subject) and M-Adj/adv[10] (adjective corrected to adverb form, as in Figure 1) appear to cause issues at the parsing stage, especially when corrections involve complex rephrasings and/or lexical changes. Morphological errors involving nonexistent word forms are also often handled incorrectly. An example of that is the Swedish L2 sentence *Kommunikationen hade dittills skett via brev, och brevutdelning fick man fem {gånger → gångar*} om dagen* ("Communication had until then taken place by mail, and letters were delivered five times a day"), where *gångar* is an incorrect plural form of the noun *gång*, corrected to *gånger*. In such cases, the morphological analysis of L2 is identical to that of the L1 and the only usable patterns are those preserving lexical information, for which finding treebank matches is less likely.

5 Conclusions and future work

We presented a novel approach for extracting morphosyntactical error patterns from L1-L2 parallel UD treebanks and put it to the test through an example retrieval task. While performed on datasets for linguistic acceptability judgments rather than authentic learner data, our preliminary evaluation gave promising results and provided helpful insights for the further development of the tool.

Future work on the extraction method itself will focus on handling nonexistent word forms and dealing with the complexity of actual L2 data. Real-world L2 texts come with two main challenges: handling non-morphosyntactical errors, such as spelling mistakes and incorrect lexical choices, and isolating each of the grammatical errors occurring in the same sentence. We mentioned that our system extracts patterns of different sizes and at varying degrees of simplification, whose usefulness depends on the use case. This drives us to also investigate pattern selection and ranking. The latter, together with a more user-friendly interface, could contribute to the improvement the example retrieval pipeline to better suit the learners' needs. Further improvements will require addressing the L2 parsing issues identified through the our preliminary evaluation, for instance by fine-tuning a UDPipe model on L2 data, and possibly intervening on the alignment step.

[9] See Appendix B for a similar example, where the same sentence matches two patterns of different sizes.

[10] Even though SWELL classifies this as a morphological error, it is syntactical from a UD perspective.

References

Mio Arai, Masahiro Kaneko, and Mamoru Komachi. 2019. Grammatical-error-aware incorrect example retrieval system for learners of Japanese as a second language. In *Proceedings of the Fourteenth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 296–305, Florence, Italy. Association for Computational Linguistics.

Christopher Bryant, Mariano Felice, and Ted Briscoe. 2017. Automatic annotation and evaluation of error types for grammatical error correction. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 793–805, Vancouver, Canada. Association for Computational Linguistics.

Leshem Choshen, Dmitry Nikolaev, Yevgeni Berzak, and Omri Abend. 2020. Classifying syntactic errors in learner language. In *Proceedings of the 24th Conference on Computational Natural Language Learning*, pages 97–107, Online. Association for Computational Linguistics.

Leshem Choshen, Matanel Oren, Dmitry Nikolaev, and Omri Abend. 2021. SERRANT: a syntactic classifier for english grammatical error types. *arXiv preprint arXiv:2104.02310*.

Mariano Felice, Christopher Bryant, and Ted Briscoe. 2016. Automatic extraction of learner errors in ESL sentences using linguistically enhanced alignments. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 825–835, Osaka, Japan. The COLING 2016 Organizing Committee.

Prasanth Kolachina and Aarnte Ranta. 2016. From abstract syntax to Universal Dependencies. *Linguistic Issues in Language Technology*, 13.

John Lee, Keying Li, and Herman Leung. 2017. L1-L2 parallel dependency treebank as learner corpus. In *Proceedings of the 15th International Conference on Parsing Technologies*, pages 44–49, Pisa, Italy. Association for Computational Linguistics.

Arianna Masciolini. 2023. A query engine for L1-L2 parallel dependency treebanks. In *Proceedings of the 24th Nordic Conference on Computational Linguistics (NoDaLiDa)*, pages 574–587, Tórshavn, Faroe Islands. University of Tartu Library.

Arianna Masciolini and Aarne Ranta. 2021. Grammar-based concept alignment for domain-specific Machine Translation. In *Proceedings of the Seventh International Workshop on Controlled Natural Language (CNL 2020/21)*, Amsterdam, Netherlands. Special Interest Group on Controlled Natural Language.

Joakim Nivre, Marie-Catherine de Marneffe, Filip Ginter, Jan Hajič, Christopher D. Manning, Sampo Pyysalo, Sebastian Schuster, Francis Tyers, and Daniel Zeman. 2020. Universal Dependencies v2: An evergrowing multilingual treebank collection. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 4034–4043, Marseille, France. European Language Resources Association.

Aarne Ranta and Prasanth Kolachina. 2017. From Universal Dependencies to abstract syntax. In *Proceedings of the NoDaLiDa 2017 Workshop on Universal Dependencies (UDW 2017)*, pages 107–116, Gothenburg, Sweden. Association for Computational Linguistics.

Lisa Rudebeck and Gunlög Sundberg. 2021. SweLL correction annotation guidelines. In *The SweLL guideline series nr 4*, Gothenburg, Sweden. Institutionen för svenska, Göteborgs Universitet.

Milan Straka. 2018. UDPipe 2.0 prototype at CoNLL 2018 UD shared task. In *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, pages 197–207, Brussels, Belgium. Association for Computational Linguistics.

Elena Volodina, Lena Granstedt, Arild Matsson, Beáta Megyesi, Ildikó Pilán, Julia Prentice, Dan Rosén, Lisa Rudebeck, Carl-Johan Schenström, Gunlög Sundberg, et al. 2019. The SweLL language learner corpus: From design to annotation. *Northern European Journal of Language Technology*, 6:67–104.

Elena Volodina, Yousuf Ali Mohammed, and Julia Klezl. 2021. DaLAJ-a dataset for linguistic acceptability judgments for Swedish: Format, baseline, sharing. *arXiv preprint arXiv:2105.06681*.

Alex Warstadt, Alicia Parrish, Haokun Liu, Anhad Mohananey, Wei Peng, Sheng-Fu Wang, and Samuel R. Bowman. 2020. BLiMP: The benchmark of linguistic minimal pairs for English. *Transactions of the Association for Computational Linguistics*, 8:377–392.

A Application of simplification strategies

Input correct-incorrect sentence pair: ⟨*I write slowly*, *I write slow*⟩.

0. largest complete extracted error pattern:

```
TREE
    (AND [
        FORM "write",
        LEMMA "write",
        POS "VERB",
        XPOS "VBP",
        FEATS "Mood=Ind|Number=Sing|Person=1|Tense=Pres|VerbForm=Fin",
        DEPREL "root"])
    [AND [
        FORM "I",
        LEMMA "I",
        POS "PRON",
        XPOS "PRP",
        FEATS "Case=Nom|Number=Sing|Person=1|PronType=Prs",
        DEPREL "nsubj"],
     AND [
        FORM {"slowly" → "slow"},
        LEMMA {"slowly" → "slow"},
        POS {"ADV" → "ADJ"},
        XPOS {"RB" → "JJ"},
        FEATS "_",
        DEPREL {"advmod" → "amod"}]]
```

1. filtering by CoNNL-U field, keeping only morphosyntax-related fields (`UPOS`, `FEATS` and `DEPREL`):

```
TREE
    (AND [
        POS "VERB",
        FEATS "Mood=Ind|Number=Sing|Person=1|Tense=Pres|VerbForm=Fin",
        DEPREL "root"])
    [AND [
        POS "PRON",
        FEATS "Case=Nom|Number=Sing|Person=1|PronType=Prs",
        DEPREL "nsubj"],
     AND [
        POS {"ADV" → "ADJ"},
        FEATS "_",
        DEPREL {"advmod" → "amod"}]]
```

2. removal of fields whose values are identical everywhere in both the L1 and L2 component:

```
TREE
    (AND [POS "VERB", DEPREL "root"])
    [AND [POS "PRON",DEPREL "nsubj"],
     AND [POS {"ADV" → "ADJ"},DEPREL {"advmod" → "amod"}]]
```

3. elimination of identical subpatterns:

```
TREE (TRUE) [TRUE, AND [POS {"ADV" → "ADJ"},DEPREL {"advmod" → "amod"}]]
```

4. monolingual single-pattern simplifications:

```
AND [POS {"ADV" → "ADJ"},DEPREL {"advmod" → "amod"}]
```

B Example program output

Input correct-incorrect sentence pair:[11] ⟨*jag skriver långsamt*, *jag skriver långsam*⟩.

Sentence 391

L1 sentence	L2 sentence
För det andra kommer studenterna ibland så **tidigt** så de måste vänta i en korridor istället för att vänta på ett café och dricka kaffe eller te .	För det andra kommer studenterna ibland så **tidig** så de måste vänta i en korridor istället för att vänta på ett café och dricka kaffe eller .

Sentence 395

L1 sentence	L2 sentence
När man inte har någon bil , får man promenera till jobbet eller ta bussen ; Det går inte så **snabbt** , och man måste planera lite mer , men det är naturligt för oss .	När man inte har någon bil , får man promenera till jobbet eller ta bussen ; Det går inte så **snabb** , och man måste planera lite mer , men det är naturligt för oss .

Sentence 684

L1 sentence	L2 sentence
Och just nu känns vårt liv **jättebra** . Och just nu **känns** vårt liv **jättebra** .	Och just nu känns våras liv **jättebra** . Och just nu **känns** våras liv **jättebra** .

Sentence 459

L1 sentence	L2 sentence
På senare år har engelskan kommit att få en allt starkare ställning **internationellt** och också i Sverige . På senare år har engelskan kommit att **få** en allt starkare ställning **internationellt** och också i Sverige .	På senare år har engelskan kommit att få en allt starkare ställning **internationell** och också i Sverige . På senare år har engelskan kommit att **få** en allt starkare ställning **internationell** och också i Sverige .

Sentence 436

L1 sentence	L2 sentence
Jag är väldigt glad över det eftersom jag tycker att det finns för många människor , **speciellt** barn , som ser kläder som en statussymbol och köper dem även om de har inte tillräckligt med pengar .	Jag är väldigt glad över det eftersom jag tycker att det finns för många människor , **speciell** barn , som ser kläder som en statussymbol och köper dem även om de har inte tillräckligt med pengar .

[11] Results obtained on the DALAJ treebank with the latest version of the interactive example retrieval pipeline (`example` command of L2-UD, run with the `-markdown` option), with commit SHA `9a1ec851313a4c3176826c77aa677e94158c3519`. As it is to be expected, some sentences match several of the extracted patterns. While seemingly identical matches have been manually removed for the sake of compactness, highlighting clearly shows that sentences like 459 match not only the single-token `POS {"ADV"→"ADJ"}` pattern, but also the more specific `TREE_ (POS "VERB") [POS {"ADV"→"ADJ"}]` pattern and could therefore be ranked higher.

Sentence 1017

L1 sentence	L2 sentence
Men i Sverige går det **bättre** för bönderna !	Men i Sverige går det **bästa** för bönderna !

Sentence 437

L1 sentence	L2 sentence
Om man skulle välja att gå emot normen så skulle det leda till utanförskap , vilket är någonting jag inte tror att någon vill uppleva , och därför väljer jag att klä mig **likadant** som de andra på mitt jobb . Om man skulle välja att gå emot normen så skulle det leda till utanförskap , vilket är någonting jag inte tror att någon vill uppleva , och därför väljer jag att **klä** mig **likadant** som de andra på mitt jobb .	Om man skulle välja att gå emot normen så skulle det leda till utanförskap , vilket är någonting jag inte tror att någon vill uppleva , och därför väljer jag att klä mig **likadan** som de andra på mitt jobb . Om man skulle välja att gå emot normen så skulle det leda till utanförskap , vilket är någonting jag inte tror att någon vill uppleva , och därför väljer jag att **klä** mig **likadan** som de andra på mitt jobb .

Sentence 420

L1 sentence	L2 sentence
Det finns **säkert** en del som undrar varför de finska ungdomarna obligatoriskt ska läsa svenska i finska skolor när endast cirka sex procent av befolkningen läser svenska som modersmål . Det **finns säkert** en del som undrar varför de finska ungdomarna obligatoriskt ska läsa svenska i finska skolor när endast cirka sex procent av befolkningen läser svenska som modersmål .	Det finns **säker** en del som undrar varför de finska ungdomarna obligatoriskt ska läsa svenska i finska skolor när endast cirka sex procent av befolkningen läser svenska som modersmål . Det **finns säker** en del som undrar varför de finska ungdomarna obligatoriskt ska läsa svenska i finska skolor när endast cirka sex procent av befolkningen läser svenska som modersmål .

Sentence 407

L1 sentence	L2 sentence
Andra punkten : Vi behöver biblioteket för att där finns böcker på olika språk , **specifikt** mitt modersmål .	Andra punkten : Vi behöver biblioteket för att där finns böcker på olika språk , **specifik** mitt modersmål .

Sentence 392

L1 sentence	L2 sentence
Det är viktigt för mig när jag behöver ta det lite **lugnt** och göra mina läxor , och det är viktigt för mig att prata svenska med en svensk person och lära mig många nya ord .	Det är viktigt för mig när jag behöver ta det lite **lugna** och göra mina läxor , och det är viktigt för mig att prata svenska med en svensk person och lära mig många nya ord .

Sentence 425

L1 sentence	L2 sentence
Historier som från början bara var **muntligt** berättade tar idag alla tänkbara former och förekommer som musik , teater , romaner , serier , filmer och spel .	Historier som från början bara var **muntlig** berättade tar idag alla tänkbara former och förekommer som musik , teater , romaner , serier , filmer och spel .

Sentence 429

L1 sentence	L2 sentence
I boken ” Stjärnlösa nätter ” så ser man tydligt hur en hatkärlek kan påverka en människas liv både **negativt** och positivt .	I boken ” Stjärnlösa nätter ” så ser man tydligt hur en hatkärlek kan påverka en människas liv både **negativ** och positivt .

Sentence 984

L1 sentence	L2 sentence
Där sitter jag med min familj och äter , sjunger , dansar , skrattar , leker och studerar . . . I hemmet kommer jag jättenära min son och jag kan lära honom mycket om livet och **hur** han kan bli bra person .	Där sitter jag med min familj och äter , sjunger , dansar , skrattar , leker och studerar . . . I hemmet kommer jag **jättenärmare** min son och jag kan lära honom mycket om livet och hur han kan bli bra person .

Sentence 401

L1 sentence	L2 sentence
Jag tycker att buss är bättre än bil eftersom det är lättare att använda buss än bil , för alla människor , **särskilt** de fattiga , kan använda buss som de vill .	Jag tycker att buss är bättre än bil eftersom det är lättare att använda buss än bil , för alla människor , **särskild** de fattiga , kan använda buss som de vill .

Sentence 457

L1 sentence	L2 sentence
Jag lärde mig att om saker inte går bra för dig ska du vara modig och ta det **lugnt** , det kommer att bli bättre , ge bara aldrig upp ! Jag lärde mig att om saker inte går bra för dig ska du vara modig och **ta** det **lugnt** , det kommer att bli bättre , ge bara aldrig upp !	Jag lärde mig att om saker inte går bra för dig ska du vara modig och ta det **lugn** , det kommer att bli bättre , ge bara aldrig upp ! Jag lärde mig att om saker inte går bra för dig ska du vara modig och **ta** det **lugn** , det kommer att bli bättre , ge bara aldrig upp !

Sentence 442

L1 sentence	L2 sentence
Det finns olika sätt som man kan använda eller utrycka sig på för att kunna kommunicera med varandra , till exempel skrivet eller **muntligt** med hjälp av ord på en mängd olika språk .	Det finns olika sätt som man kan använda eller utrycka sig på för att kunna kommunicera med varandra , till exempel skrivet eller **muntlig** med hjälp av ord på en mängd olika språk .

Sentence 421

L1 sentence	L2 sentence
Detta leder till motstånd från landets folk som ser **negativt** på regeringens maktfullkomliga metod . Detta leder till motstånd från landets folk som **ser negativt** på regeringens maktfullkomliga metod .	Detta leder till motstånd från landets folk som ser **negativ** på regeringens maktfullkomliga metod . Detta leder till motstånd från landets folk som **ser negativ** på regeringens maktfullkomliga metod .

Sentence 431

L1 sentence	L2 sentence
Historier som från början bara var **muntligt** berättade tar idag alla tänkbara former och förekommer som musik , teater , poesi , romaner , serier , filmer och spel .	Historier som från början bara var **muntliga** berättade tar idag alla tänkbara former och förekommer som musik , teater , poesi , romaner , serier , filmer och spel .

Sentence 458

L1 sentence	L2 sentence
Det är inte så lätt att svara **snabbt** . Det är inte så lätt att **svara snabbt** .	Det är inte så lätt att svara **snabb** . Det är inte så lätt att **svara snabb** .

Sentence 451

L1 sentence	L2 sentence
Mitt råd är att du måste ta det **lugnt** och fokusera , till exempel klä på dig fina kläder , det betyder inte smustiga kläder , eller du kan använda parfym , men inte så mycket . Mitt råd är att du måste **ta** det **lugnt** och fokusera , till exempel klä på dig fina kläder , det betyder inte smustiga kläder , eller du kan använda parfym , men inte så mycket .	Mitt råd är att du måste ta det **lugn** och fokusera , till exempel klä på dig fina kläder , det betyder inte smustiga kläder , eller du kan använda parfym , men inte så mycket . Mitt råd är att du måste **ta** det **lugn** och fokusera , till exempel klä på dig fina kläder , det betyder inte smustiga kläder , eller du kan använda parfym , men inte så mycket .

Sentence 466

L1 sentence	L2 sentence
Jag personligen lägger inte **medvetet** så stor vikt vid kläder , kanske för att den miljö som jag lever i eller de människor som jag umgås med inte ser kläder som något betydelsefullt . Jag personligen **lägger** inte **medvetet** så stor vikt vid kläder , kanske för att den miljö som jag lever i eller de människor som jag umgås med inte ser kläder som något betydelsefullt .	Jag personligen lägger inte **medveten** så stor vikt vid kläder , kanske för att den miljö som jag lever i eller de människor som jag umgås med inte ser kläder som något betydelsefullt . Jag personligen **lägger** inte **medveten** så stor vikt vid kläder , kanske för att den miljö som jag lever i eller de människor som jag umgås med inte ser kläder som något betydelsefullt .

Sentence 462

L1 sentence	L2 sentence
Alla mina dagar gick så **dåligt** .	Alla mina dagar gick så **dålig** .

Sentence 461

L1 sentence	L2 sentence
Sammanfattat har jag en föränderlig relation till kläder , men det viktigaste är att de möjliggör allt jag vill uppleva , från bergsvandring till fest . **Sammanfattat har** jag en föränderlig relation till kläder , men det viktigaste är att de möjliggör allt jag vill uppleva , från bergsvandring till fest .	**Sammanfattad** har jag en föränderlig relation till kläder , men det viktigaste är att de möjliggör allt jag vill uppleva , från bergsvandring till fest . **Sammanfattad har** jag en föränderlig relation till kläder , men det viktigaste är att de möjliggör allt jag vill uppleva , från bergsvandring till fest .

Sentence 390

L1 sentence	L2 sentence
Dessutom är det **troligen** kö då alla vill ha rast och kaffe samtidigt .	Dessutom är det **troliget** kö då alla vill ha rast och kaffe samtidigt .

Sentence 387

L1 sentence	L2 sentence
Det var ganska svårt **först** men jag är van och lärde mig själv hur man bor och anpassar sig i ett nytt land .	Det var ganska svårt **första** men jag är van och lärde mig själv hur man bor och anpassar sig i ett nytt land .

Sentence 469

L1 sentence	L2 sentence
Tänk **positivt** istället så kommer du att hitta många betydelsefulla saker inom din familj . **Tänk positivt** istället så kommer du att hitta många betydelsefulla saker inom din familj .	Tänk **positiv** istället så kommer du att hitta många betydelsefulla saker inom din familj . **Tänk positiv** istället så kommer du att hitta många betydelsefulla saker inom din familj .

Sentence 467

L1 sentence	L2 sentence
Detta kan dock skapa svårigheter med att kunna förbereda och undervisa ungdomar **tillräckligt** . Detta kan dock skapa svårigheter med att kunna **förbereda** och undervisa ungdomar **tillräckligt** .	Detta kan dock skapa svårigheter med att kunna förbereda och undervisa ungdomar **tillräckliga** . Detta kan dock skapa svårigheter med att kunna **förbereda** och undervisa ungdomar **tillräckliga** .

Sentence 410

L1 sentence	L2 sentence
Efter några år visade inspektörerna rapporter om att det nog fanns lite kokain i coca cola , men tyvärr ville de inte kommunicera detta **offentligt** . Efter några år visade inspektörerna rapporter om att det nog fanns lite kokain i coca cola , men tyvärr ville de inte **kommunicera** detta **offentligt** .	Efter några år visade inspektörerna rapporter om att det nog fanns lite kokain i coca cola , men tyvärr ville de inte kommunicera detta **offentlig** . Efter några år visade inspektörerna rapporter om att det nog fanns lite kokain i coca cola , men tyvärr ville de inte **kommunicera** detta **offentlig** .

Sentence 375

L1 sentence	L2 sentence
Hon lär ut svenska mycket **snällt** och fint .	Hon lär ut svenska mycket **snäll** och fint .

Sentence 463

L1 sentence	L2 sentence
Jag hoppas kunna lära mig **snabbt** och börja söka jobb .	Jag hoppas kunna lära mig **snabb** och börja söka jobb .
Jag hoppas kunna **lära** mig **snabbt** och börja söka jobb .	Jag hoppas kunna **lära** mig **snabb** och börja söka jobb .