

## **Categorizing Teachers' Corrections: Guidelines for Annotating the Šolar Corpus**

Špela ARHAR HOLDT, Polona LAVRIČ, Rebeka ROBLEK, and Teja GOLI

**V1.1 (12/8/2022)**

### **Dokumentacija verzij**

V1.0	16/12/2018	The project <i>Nadgradnja korpusa Šolar</i> ('Upgrading the Šolar Corpus'), contract 3340-15-141006, under the call <i>Digitalizacija na področju kulture 2015-2018</i> , was funded by the Ministry of Culture of the Republic of Slovenia. The results were supported by the <i>Center za uporabno jezikoslovje</i> ('Centre for Applied Linguistics') programme, which is funded by the ARRS from the state budget. V1.0 is followed by the Šolar 2.0 corpus, available at <a href="http://hdl.handle.net/11356/1214">http://hdl.handle.net/11356/1214</a> .
V1.1	12/8/2022	The project <i>Development of Slovene in the Digital Environment</i> was Financed by the Ministry of Culture and the European Regional Development Fund. The project <i>Empirical foundations for digitally-supported development of writing skills</i> (J7-3159) and the programme <i>Language resources and technologies for Slovene language</i> (P6-0411) were financed by the ARRS from the state budget. V1.1 of the guidelines includes some minor improvements in the structure of certain labels and adds chapter 5.5. V1.1 is followed by the Šolar 3.0 corpus, available at <a href="http://hdl.handle.net/11356/1589">http://hdl.handle.net/11356/1589</a> .

## Kazalo vsebine

<b>1 SPELLING .....</b>	<b>6</b>
1.1 Vowels.....	6
1.1.1 Superfluous vowel = Č/VOK/odveč .....	6
1.1.2 Omitted vowel = Č/VOK/izpust.....	7
1.1.3 AO substitution = Č/VOK/menjava-ao.....	7
1.1.4 EI substitution = Č/VOK/menjava-ei .....	8
1.1.5 UO substitution = Č/VOK/menjava-uo.....	8
1.1.6 Other vowel substitutions = Č/VOK/menjava-drugo .....	8
1.2 Consonants .....	9
1.2.1 Superfluous consonant = Č/KONZ/odveč.....	9
1.2.2 Omitted consonant = Č/KONZ/izpust.....	10
1.2.3 Substitution SZ = Č/KONZ/menjava-sz .....	10
1.2.4 Substitution TD = Č/KONZ/menjava-td .....	10
1.2.5 Substitution KGH = Č/KONZ/menjava-kgh.....	11
1.2.6 Substitution MN = Č/KONZ/menjava-mn .....	11
1.2.7 Substitution ŠŽ = Č/KONZ/menjava-šž .....	11
1.2.8 Substitution of DIACRITIC = Č/KONZ/menjava-strešice .....	11
1.2.9 Other consonant substitution = Č/KONZ/menjava-drugo .....	12
1.3 Labio-velar approximant w .....	12
1.3.1 Word-initially = Č/W/začetek .....	12
1.3.2 Word-medial = Č/W/sredina.....	13
1.3.3 Word-final = Č/W/konec .....	13
1.3.4 Prepositional V = Č/W/v.....	14
1.4 Letter clusters .....	14
1.4.1 A syllable is missing or is superfluous = Č/SKLOP/zlog .....	14
1.4.2 Cluster LJ = Č/SKLOP/lj .....	15
1.4.3 Cluster NJ = Č/SKLOP/nj .....	16
1.4.4 Cluster IJ = Č/SKLOP/ij.....	16
1.4.5 Doubled letters = Č/SKLOP/podvojene .....	17
1.4.6 Metathesis = Č/SKLOP/premet.....	18
1.5 Variable (allophonic) prepositions .....	18
1.5.1 Preposition s/z = Č/PREDL/sz.....	18
1.5.2 Preposition k/h = Č/PREDL/kh.....	19
<b>2 MORPHOLOGY .....</b>	<b>19</b>
2.1 Categorical corrections .....	19
2.1.1 Case: genitive-accusative = O/KAT/sklon-rt .....	19
2.1.2 Case: dative-locative = O/KAT/sklon-dm.....	20
2.1.3 Case: locative-instrumental = O/KAT/sklon-mo .....	20
2.1.4 Other case substitutions = O/KAT/sklon-drugo .....	21
2.1.5 Number: singular-plural = O/KAT/število-em .....	22
2.1.6 Number: dual-plural = O/KAT/število-dm.....	22

2.1.7	Number: single-dual = O/KAT/število-ed .....	23
2.1.8	Gender = O/KAT/spol .....	23
2.1.9	Aspect = O/KAT/vid .....	24
2.1.10	Tense = O/KAT/čas .....	24
2.1.11	Person = O/KAT/oseba .....	25
2.1.12	Reduced infinitive = O/KAT/nedoločnik-kratki .....	25
2.1.13	Infinitive and supine = O/KAT/nedoločnik-namenilnik .....	25
2.1.14	Infinitive and finite verb = O/KAT/nedoločnik-osebna .....	26
2.1.15	Reflexivity = O/KAT/povratnost .....	26
2.1.16	Mood = O/KAT/naklon .....	27
2.1.17	Voice = O/KAT/način .....	27
2.1.18	Pronomial form = O/KAT/oblika-zaimka .....	28
2.1.19	Definiteness = O/KAT/določnost .....	28
2.1.20	Comparison = O/KAT/stopnjevanje .....	28
2.2	Paradigmatic Corrections .....	29
2.2.1	Verbal root = O/PAR/glagolska-osnova .....	29
2.2.2	Verbal ending = O/PAR/glagolska-končnica .....	30
2.2.3	Non-verbal root = O/PAR/neglagolska-osnova .....	30
2.2.4	Non-verbal ending = O/PAR/neglagolska-končnica .....	31
2.2.5	Epenthetic vowel = O/PAR/neobstojni-vokal .....	32
2.2.6	Umlaut and cč = O/PAR/preglas-in-cč .....	32
2.3	Additional Annotation .....	33
2.3.1	Morphological variants = O/DOD/variante .....	33
2.3.2	Mati-hči = O/DOD/besede-mati-hči .....	34
2.3.3	Otrok = O/DOD/besede-otrok .....	34
<b>3</b>	<b>VOCABULARY .....</b>	<b>34</b>
3.1	Nouns .....	35
3.1.1	Erroneous proper noun = B/SAM/napačno-lastno .....	35
3.1.2	Proper and common name = B/SAM/lastno-občno .....	35
3.1.3	Common vocabulary = B/SAM/občno-besedišče .....	36
3.2	Verbs .....	36
3.2.1	Verbal prefixes = B/GLAG/predpona .....	36
3.2.2	Substitution moči-morati = B/GLAG/moči-morati .....	37
3.2.3	Other substitutions of modal verbs = B/GLAG/naklonski .....	37
3.2.4	Other substitutions of verbs = B/GLAG/drugo .....	38
3.3	Pronoun .....	39
3.3.1	Reflexive possessive = B/ZAIM/povratna-svojlnost .....	39
3.3.2	Substitution ki – kateri = B/ZAIM/ki-kateri .....	39
3.3.3	Other problems with relative pronouns = B/ZAIM/oziralni .....	39
3.3.4	Substitution of negative pronouns = B/ZAIM/noben .....	40
3.3.5	Other pronomial substitutions = B/ZAIM/drugo .....	40
3.4	Preposition .....	40
3.4.1	Prepositions in verbal phrases = B/PRED/glag-zveze .....	40
3.4.2	Prepositions into non-verbal phrases = B/PRED/neglag-zveze .....	41
3.4.3	Locative doublets = B/PRED/lokacijske-dvojnice .....	42

3.4.4	Other substitutions of prepositions = B/PRED/drugo .....	42
3.5	Conjunction .....	43
3.5.1	Substitutions of in-pa-ter = B/VEZ/in-pa-ter .....	43
3.5.2	Coordinating adversative conjunction = B/VEZ/protivni .....	43
3.5.3	Change to subordination = B/VEZ/sprememba-odnosa .....	43
3.5.4	Other substitutions of conjunctions B/VEZ/drugo .....	44
3.6	Adjective = B/PRID/drugo.....	44
3.7	Adverb = B/PRISL/drugo.....	45
3.8	Other parts of speech = B/OST/drugo.....	45
3.9	Substitutions beyond the confines of part of speech .....	46
3.9.1	Lexical word or phrase changed into pronoun = B/MEN/polnopomenska-v-zaimek	46
3.9.2	Pronoun to a lexical word or phrase = B/MEN/zaimek-v-polnopomensko	46
3.9.3	Substitution of conjunction and pronoun = B/MEN/veznik-zaimek .....	46
3.9.4	Word family = B/MEN/besedna-družina.....	47
3.9.5	Noun and phrase = B/MEN/samostalnik-bz.....	47
3.9.6	Verb and phrase = B/MEN/glagol-bz.....	48
3.9.7	Adverb/adjective and phrase = B/MEN/prislov-pridevnik-bz .....	48
3.9.8	Other types of substitutions =B/MEN/drugo .....	49
3.10	Additional Annotation .....	49
3.10.1	Stylistically marked vocabulary = B/DOD/zaznamovano .....	49
<b>4</b>	<b>SYNTAX.....</b>	<b>50</b>
4.1	Word order .....	50
4.1.1	Constituent order: predicate-subject = S/BR/povedek-osebek.....	50
4.1.2	Constituent order: povedek-predmet = S/BR/povedek-predmet .....	51
4.1.3	Order: sentence-adverbial determiner = S/BR/povedek-prislovno-določilo	51
4.1.4	Order: particle = S/BR/členek .....	52
4.1.5	Order within clausal constituents = S/BR/znotraj-stavčnega-člena .....	52
4.1.6	Clitic string: order of clitics = S/BR/naslonski-niz-znotraj .....	52
4.1.7	Clitic string: independent-subordinate = S/BR/naslonski-niz-prirednost-podrednost	52
4.1.8	Other changes to word order = S/BR/drugo .....	53
4.2	Omitted constituents .....	54
4.2.1	Noun: common noun = S/IZPUST/samostalnik-občno-ime.....	54
4.2.2	Noun: proper noun = S/IZPUST/samostalnik-lastno-ime .....	54
4.2.3	The verb biti = S/IZPUST/glagol-biti .....	55
4.2.4	Other omitted verbs = S/IZPUST/glagol-drugo.....	55
4.2.5	The word pa = S/IZPUST/veznik-pa .....	56
4.2.6	Other omitted conjunctions = S/IZPUST/veznik-drugo.....	56
4.2.7	Repeated prepositions = S/IZPUST/predlog-ponovljen.....	56
4.2.8	Other omitted prepositions = S/IZPUST/predlog-drugo .....	57
4.2.9	Personal pronoun = S/IZPUST/zaimek-osebni .....	57
4.2.10	Other omitted pronouns = S/IZPUST/zaimek-drugo.....	57
4.2.11	Adjective = S/IZPUST/pridevnik.....	57
4.2.12	Adverb = S/IZPUST/prislov .....	58

4.2.13	Particle = S/IZPUST/členek .....	58
4.2.14	Sentence = S/IZPUST/stavek .....	58
4.3	Superfluous constituents .....	59
4.3.1	Literal repetition = S/ODVEČ/ponavljanje .....	59
4.3.2	Noun: common noun = S/ODVEČ/samostalnik-občno-ime .....	60
4.3.3	Noun: proper noun = S/ODVEČ/samostalnik-lastno-ime .....	60
4.3.4	The verb biti = S/ODVEČ/glagol-biti .....	60
4.3.5	Other superfluous verb = S/ODVEČ/glagol-drugo .....	61
4.3.6	The word pa with another conjunction = S/ODVEČ/veznik-pa-vezniki .....	61
4.3.7	Other examples including the word pa = S/ODVEČ/veznik-pa-drugo .....	61
4.3.8	Conjunction at the beginning of a sentence = S/ODVEČ/veznik-začetek .....	62
4.3.9	Doubled conjunction = S/ODVEČ/veznik-dvojni .....	62
4.3.10	Other superfluous conjunction = S/ODVEČ/veznik-drugo .....	62
4.3.11	Preposition = S/ODVEČ/predlog .....	63
4.3.12	Personal pronoun = S/ODVEČ/zaimek-osebni .....	63
4.3.13	Demonstrative pronoun = S/ODVEČ/zaimek-kazalni .....	63
4.3.14	Possessive pronoun = S/ODVEČ/zaimek-svojilni .....	64
4.3.15	Other superfluous pronouns = S/ODVEČ/zaimek-drugo .....	64
4.3.16	Adjective = S/ODVEČ/pridevnik .....	64
4.3.17	Adverb of degree = S/ODVEČ/prislov-mera .....	65
4.3.18	Other superfluous adverbs = S/ODVEČ/prislov-drugo .....	65
4.3.19	Particle = S/ODVEČ/členek .....	65
4.3.20	Clause = S/ODVEČ/stavek .....	65
4.3.21	Sentence = S/ODVEČ/poved .....	66
4.4	Structure .....	66
4.4.1	Possessives with od = S/STR/svojina-od .....	66
4.4.2	Possessives with the genitive = S/STR/svojina-rodilnik .....	66
4.4.3	Substitution punctuation-conjunction = S/STR/ločilo-veznik .....	67
4.4.4	Merged clauses = S/STR/združevanje-stavkov .....	67
4.4.5	Separation of clauses/sentences = S/STR/deljenje-stavkov .....	68
4.4.6	Word/phrase instead of clause and vice versa = S/STR/besedna-zveza-stavek .....	69
4.4.7	Reworked clause = S/STR/preoblikovanje-stavka .....	69
4.5	Additional Annotation .....	70
4.5.1	Pleonasm = S/DOD/pleonazem .....	70
4.5.2	Superfluous content = S/DOD/vsebina-drugo .....	70
4.5.3	Erroneous content = S/DOD/vsebina-napake .....	71
4.5.4	Semantically null = S/DOD/pomensko-prazni .....	71
<b>5</b>	<b>ORTHOGRAPHY .....</b>	<b>72</b>
5.1	Capital/lowercase letters .....	72
5.1.1	Adjectives ending in -ski = Z/MV/pridevnik-ski .....	72
5.1.2	Other adjectives = Z/MV/pridevnik-drugo .....	72
5.1.3	Common noun = Z/MV/občna-imena .....	72
5.1.4	Personal name with lowercase letter = Z/MV/zemljepisna-imena .....	73
5.1.5	Nationality with lowercase letter = Z/MV/narodnost .....	73
5.1.6	Geographical name with lowercase letter = Z/MV/zemljepisna-imena .....	73
5.1.7	Proper nouns with lowercase letter = Z/MV/stvarna-imena .....	73
5.1.8	Direct speech = Z/MV/premi-govor .....	74

5.1.9	Sentence initial = Z/MV/začetek-povedi .....	74
5.1.10	Hypercorrection following a period = Z/MV/hiperkorekcija-ločila .....	74
5.1.11	Other problems with initial letters = Z/MV/drugo .....	74
5.2	Together/separate .....	75
5.2.1	Verb together = Z/SN/skupaj- <i>glagol</i> .....	75
5.2.2	Preposition together = Z/SN/skupaj- <i>predlog</i> .....	75
5.2.3	Preposition separate = Z/SN/narazen- <i>predlog</i> .....	75
5.2.4	Adverb together = Z/SN/skupaj- <i>prislov</i> .....	76
5.2.5	Adverb separate = Z/SN/narazen- <i>prislov</i> .....	76
5.2.6	Adjective separate = Z/SN/narazen- <i>pridevnik</i> .....	76
5.2.7	Other separate = Z/SN/narazen- <i>drugo</i> .....	76
5.2.8	Other together = Z/SN/skupaj- <i>drugo</i> .....	77
5.3	Abbreviations = Z/KR/drugo .....	77
5.4	Numbers = Z/ŠTEV/drugo .....	77
5.5	Punctuation.....	78
5.5.1	Unclassified punctuation corrections = Z/LOČ/nerazvrščeno .....	78
5.5.2	Comma before subordinate clauses = Z/LOČ/vzorec- <i>vejica-stavki</i> .....	78
5.5.3	Comma between parts-of-speech = Z/LOČ/vzorec- <i>vejica-stavčni-členi</i> ...	78
5.5.4	Comma and multi-word conjunctions = Z/LOČ/vzorec- <i>vejica-vezniki</i> .....	79
5.5.5	Comma and comparative structures = Z/LOČ/vzorec- <i>vejica-kot</i> .....	79
5.5.6	Comma and appositions etc. = Z/LOČ/vzorec- <i>vejica-pristavki</i> .....	79
5.5.7	Comma and inserted subordinate clauses = Z/LOČ/vzorec- <i>vejica-vrinjen-odvisnik</i>	79
5.5.8	Comma and coordinate phrases = Z/LOČ/vzorec- <i>vejica-piredja-zvez</i> .....	79
5.5.9	Comma and coordinate clauses = iZ/LOČ/vzorec- <i>vejica-piredja-odvisnikov</i>	79
5.5.10	Comma in adjective strings = Z/LOČ/vzorec- <i>vejica-pridevniški-niz</i> .....	79
5.5.11	Comma and predicate ellipsis = Z/LOČ/vzorec- <i>vejica-elipsa-povedka</i> .....	80
<b>6</b>	<b>RELATED CORRECTIONS .....</b>	<b>80</b>
6.1	P/OBL/drugo .....	80
6.2	P/SKLA/osebik .....	81
6.3	P/SKLA/drugo.....	81
6.4	P/ZAP/mala-velika .....	81
<b>7</b>	<b>ILLEGIBLE AND DUBIOUS EXAMPLES.....</b>	<b>82</b>
7.1	Illegible examples = N// <i>necitljivo</i> .....	82
7.2	Dubious examples = N// <i>preveri</i> .....	82
<b>8</b>	<b>DECIDING BETWEEN FRINGE CASES .....</b>	<b>82</b>

## 1 SPELLING

Here we categorize teacher's spelling corrections as they relate to how students represented a phoneme or group thereof. In distinguishing spelling corrections from morphology corrections it is crucial to identify where within the word the problem occurred. If the correction targets inflection, i.e. within the word's morphological endings or within root transformation as a result of inflection, the correction is thus categorized as morphological. The same holds true for corrections referencing morphological categories such as gender, case, number, etc. Ambiguities also arise in distinguishing between the level of spelling and the level of vocabulary. Corrections that change one word into another (e.g. *preigral – priigral*) are classified as corrections to vocabulary under the condition that such an interpretation is well-founded given the context of the passage. More on the problems surrounding classification is found in Chapter 8.

### 1.1 Vowels

This category contains corrections to vowels at the level of spelling.

#### 1.1.1 Superfluous vowel = Č/VOK/odveč

These examples show where teachers have corrected a superfluous vowel within a word. The majority of examples concern erroneous transcriptions of the epenthetic unstressed neutral vowel *e* (hereinafter: schwa). In accordance with the general guidelines for annotation, problems related to roots or base forms are regarded as spelling mistakes (while those related to inflection are categorized as morphological).

##### Č/VOK/odveč:

- Da je to svet **polen | poln** laži in prevar, v katerem je pomemben bliš in zunanji videz.
- In vse to privede do takega zaključka, ki pa se mi ne zdi **smiselen | smiseln**.
- Prometeju to spet ni bilo všeč in se je spet **uprel | uprl**.

Whenever a teacher corrects a superfluous letter cluster otherwise containing a vowel, such a correction is annotated as a superfluous letter cluster.

##### Č/SKLOP/zlog:

- Vendar **menimin | menim**, da si zaradi ljubezni nebi nikoli stregla po življenju oziroma si ga vzela kakor so si ga Antigona, Hajmon in Evridika, čeprav nikoli neves, kaj si pripravljen storiti za ljubezen v določenih okoliščinah.
- Vsako leto se **zberejejo | zberejo** ljudje v Podpeci in gradijo gradove kralja Matjaža.

Superfluous vowels can occur within letter clusters that have already been identified as problematic (*-ij* or doubled letters). If this is the case, it is annotated with both appropriate tags.

##### Č/VOK/odveč in Č/SKLOP/ij:

- Želi ustvariti kompromise, da bi imel čimmanj težav, če se mu pa na tak način ne izide, ukrepa s **smrtijo | smrtjo**.
- Ko smo prišli z **ladije | ladje**, smo šli v Piran, kjer smo si lahko kupili stvari.

### 1.1.2 Omitted vowel = Č/VOK/izpust

These examples show where teachers have corrected a dropped vowel within a word. As with superfluous vowels, the problem here is often linked with the omission of schwa, otherwise often in the transcription of foreign personal nouns.

#### Č/VOK/izpust:

- Sprva ni želel postati kristjan, saj so bili že njegovi starši in prastarši pogani in je še zmeraj **vrjel | verjel** v pogansko vero.
- Napisal jo je William **Shakespear | Shakespeare**, ki je bil angleški dramatik.
- Mariji se je zdel **čudn | čuden** človek, a ga je vseeno ljubila.

Omitted vowels can occur within letter clusters that have already been identified as problematic (-ij or doubled letters). If this is the case, it is annotated with both appropriate tags.

#### Č/VOK/izpust in Č/SKLOP/ij:

- Zgodba se začne dogajati v **italjanskem | italijanskem** mestu Verona, kjer je spet izbruhnil prepir med sptima družinama Capulet in Manteg.
- V obeh delih avtorja prikazujeta nesmiselnost takratnega časa (**socjalizma | socializma**), njegove krutosti pa tudi kako nesmiselno so obravnavali tistega, ki je nekaj zagrešil.

### 1.1.3 AO substitution = Č/VOK/menjjava-ao

These examples feature teachers changing the letter *a* to *o* or vice versa.

#### Č/VOK/menjjava-ao:

- Tudi če kej naredim narode ali kaj, s čimer ona ni zadovoljna, mi pame, da je moja mama **prastitutka | prostitutka**.
- Ko pa so otroci odrasli so tudi oni živeli podobno kot Meta, saj so se prav tako v mladih letih odpravili na **pat | pot** za kruhom.
- To je moje mnenje in zgodbo o revščini nekoč in **dones | danes**.

This category also includes corrections to spelling under the influence of o-a vowel reduction. Although many such corrections address problems in the ending, they are easier to classify at the level of spelling than morphology, as in truth they do not indicate the incorrect use of a morphological category and its pertinent ending, but rather dialectal influences on orthography (similarly the even less common occurrences of o-u vowel reduction are categorized as spelling mistakes). We are generally hesitant in determining whether an example fits the criteria of o-a vowel reduction, relying on metatags and the broader context.

#### Č/VOK/menjjava-ao:

- V današnjem svetu je vse polno takih društev, ki **zgliedaja | zgliedajo** navzven lepi, urejeni.



#### 1.1.4 EI substitution = Č/VOK/menjava-ei

These examples feature teachers changing the letter *e* to *i* or vice versa.

Č/VOK/menjava-ei:

- Dobro je, da prisluhnemo drugim, ki se na stvari spoznajo, ali pa le zato, da mnenja **primirjamo | primerjamo**.
- Rojena je bila za ta ples in že od majhnih nog je **trenerala | trenirala** ta ples s soplesalcem.
- **Črtomer | Črtomir** je moral iti v vojsko.

Examples where the correction is in a word that has, in its correct form, doubled vowels receive a tag for doubled letters in addition to the tag EI.

Č/VOK/menjava-ei in Č/SKLOP/podvojene:

- Brdavs (Angleški bogataš) je vsak dan in noč **preigral | priigral** sebi milijon evrov.

#### 1.1.5 UO substitution = Č/VOK/menjava-uo

These are examples where a teacher has corrected the letter *o* into *u* or vice versa.

Č/VOK/menjava-uo:

- Vztrajati je smiselno v tistih rečeh ko veš, da ti bo nekoč **ospelo | uspelo**, ne pa v tistih ki veš da ne bodo **ospele | uspele**.
- Ko sem napisal mamici SMS mi je odgovorila, da je tudi ona vesela, ker sem dobil dobro **uceno | oceno**.
- Edina gospodarska dejavnost je tovarna **akumolatorskih | akumulatorskih** baterij.

Examples where the correction is in a word that has, in its correct form, doubled vowels receive a tag for doubled letters in addition to the tag UO.

Č/VOK/menjava-uo in Č/SKLOP/podvojene:

- Lahko pa naletimo na primer, ki nam lahko naredi veliko škode, če se z njim **suočimo | soočimo**.

#### 1.1.6 Other vowel substitutions = Č/VOK/menjava-drugo

These examples show where teachers have corrected some other improper vowel usage within a word (a catch-all category encompassing all substitutions that occurred in the corpus' initial version fewer than 35 times).

Č/VOK/menjava-drugo:

- Zamoti se le s pisanjem o **asimulantu | asimilantu**, izdajaletlju slovenstva Jerneju Jerobniku, zanimive pa se mu zdijo tudi razprave z njegovim profesorjem De Martinisom, čeprav se njuna mnenja pogosto izključujejo.
- Nevemo kako sta prišli do Pirana, ampak tam so jih videli trije **klošerji | klošarji**.
- Ne zavedajo se, kakšno škodo delajo, **kuko | kako** to vpliva na podnebje.

Examples featuring two vowels changing places are treated as metathesis.

Č/SKLOP/premet:

- **Opozavalci | Opazovalci** se ne zmenijo za doganjo na avtobusu, saj se jim smrad od ženske zdi vsakdanji.

Examples where students write a vowel instead of a consonant and vice versa are not treated as substitutions, but are annotated as dropped/superfluous consonants and dropped/superfluous vowels.

Č/VOK/odveč in Č/KONZ/izpust:

- Djuro je Bronjo celo tako ljubil, da bi zanjo ubijal in **eradel | kradel**, vendar se moramo zavedati, da je tudi to zločin.

## 1.2 Consonants

### 1.2.1 Superfluous consonant = Č/KONZ/odveč

These are examples where a teacher has corrected a superfluous consonant in a word. These problems often arise in transcribing (foreign) surnames.

Č/KONZ/odveč:

- Tujec je **eksinstencialistični | eksistencialistični** roman, po obliki in slogi pa tradicionalen.
- In tudi **bogatstvo | bogastvo** ne pomeni vedno sreče, če ni pravega odnosa.
- Po vojni je z Lukežem in **Schwartzkoblerjem | Schwarzkoblerjem** dobil blagajno polno cekinov.

The common mistake of misspelling *jaz* as *jast* is annotated both for superfluous consonants as well as a SZ substitution.

Č/KONZ/odveč in Č/KONZ/menjava-sz:

- Upam, da me nisi pozabil, ker **jast | jaz** nate nisem in nebom.

Whenever a teacher corrects a superfluous letter cluster otherwise containing a consonant, such a correction is annotated (only) as a superfluous letter cluster.

Č/SKLOP/zlog:

- Vsako leto se **zberejejo | zberejo** ljudje v Podpeci in gradijo gradove kralja Matjaža.
- Seveda, v obdobju miru bi na ista vprašanja, od istih oseb **dobilili | dobili** popolnoma drugačne odgovore.

Superfluous consonants can be corrected within letter clusters that have already been identified as problematic (*lj, nj, ij*, or doubled letters). If this is the case, it is annotated with both appropriate tags.

Č/KONZ/odveč in Č/SKLOP/podvojene:

- Roman Tujec je delo Alberta **Cammusa | Camusa**.

- Narodna univerzitetna **knjižnica | knjižnica**.

### 1.2.2 Omitted consonant = Č/KONZ/izpust

In these examples teachers add an omitted consonant (corrections often occur in the middle and at the end of a word).

Č/KONZ/izpust:

- Danes ni tako, da starši odločajo, **apak | ampak** se lahko sam ni važno bogastvo.
- Povedala mi je, da če ne bo našla snubca, da bo odšla v **samosta | samostan**.

Omitted consonants can be corrected within letter clusters that have already been identified as problematic (*lj, nj, ij*, or doubled letters). If this is the case, it is annotated with both appropriate tags.

Č/KONZ/izpust in Č/SKLOP/podvojene:

- V večerni **odaji | oddaji** je bila Davidova smrt poročana.
- Čakal sem in čakal ter pogledal **iza | izza** roba banke NLB.

Some examples seem at first to belong to the category of vocabulary choice. Our annotation choice is thus based on context and on analogy with other typical corrections at both levels.

Č/KONZ/izpust:

- Naredila jim je velik **riž | križ**, ker je sama delala cokle, je še vsakemu podtaknila pod pazduho par novih cokel.
- Meta glavna oseba, **izjema | izjemna** mati, nosilka pozitivnega.

### 1.2.3 Substitution SZ = Č/KONZ/menjava-sz

These are examples where a teacher has corrected the substitution of s and z in a word.

Č/KONZ/menjava-sz:

- Kreon je bil trmast in **vstrajen | vztrajen** – Antigona je morala umreti, čeprav bi pri tem izgubil sina.
- Nezakonskih mater in zapuščenih žensk je **danez | danes** zelo veliko in cedalje več, saj se mnogi danes ne poročajo vel ali pa sami vzgajajo otroke.
- To pa zato, ker ima značilnosti epike, da je to neko prozno besedilo in značilnosti lirike, ki pa **ispovedujejo | izpovedujejo** čustva učencev.

### 1.2.4 Substitution TD = Č/KONZ/menjava-td

These are examples where a teacher has corrected the substitution of *t* and *d* in a word.

Č/KONZ/menjava-td:

- Linhart je zelo navdušila Figarova **svadba | svatba**.
- Kljub temu, da je bila prej skupina ljudi za **truگو | drugo** stranko, se nekatere da prepričati z dobrimi argumenti.
- Hodijo v službe in jim ni treba hoditi v vojne **kod | kot** v antičnih časih.

### 1.2.5 Substitution KGH = Č/KONZ/menjava-kgh

These are examples where a teacher has corrected various consonant substitutions between *k*, *g*, and *h*.

Č/KONZ/menjava-kgh:

- Če se nekateri **gdaj | kdaj** znajdejo v stiski z denarjem še vedno je dovolj za preživetje.
- Ko pa je Romeo prišel do Julije je seveda mislil da je mrtva in je tudi tudi popil strup in umrl **polek | poleg** nje.
- Povabljal je drugega možakarja naj ostane pri njih kot **host | gost** vendar ta je odkimal in šel na pot rekoč, da tudi on ima družino, ki ga čaka.

### 1.2.6 Substitution MN = Č/KONZ/menjava-mn

These are examples where a teacher has corrected the substitution of *m* and *n* in a word.

Č/KONZ/menjava-mn:

- Vljudno vas **prosin | prosim**, da **ni | mi** udobrite ogled vaše restavracije?
- Njene pozitivne strani so to, da lahko spomočjo nje **dobino | dobimo** iformacije, vsakdanje novice.
- Oblečena sta bila drugače kakor **sno | smo** sedaj mi oblečeni.

### 1.2.7 Substitution ŠŽ = Č/KONZ/menjava-šž

These are examples where a teacher has corrected the substitution of *š* and *ž* in a word.

Č/KONZ/menjava-šž:

- Moje mnenje je tako da ni **važno | važno** bogastvo ampak ljubezen do nekoga.
- Zato pa je na svetu tudi **sovražtvo | sovraštvo** in ljubosumje.
- Njeno življenje je bilo **teško | težko**.

### 1.2.8 Substitution of DIACRITIC = Č/KONZ/menjava-strešice

These are examples where a teacher has corrected a superfluous or omitted caron.

Č/KONZ/menjava-strešice:

- Oba **cutita | čutita** ljubezen drug do drugega ampak jih ta verski spopad ločuje drug od drugega.
- Vedenje od očeta in matere je zelo veselo, jočejo se zaradi veselja ker je končno prišel ta trenutek ko sta **žopet | zopet** oba dva skupaj.
- Toda Martin se je odmaknil in ga začaral v **kokos | kokoš**.

Problems in substituting the letter *c* for *s* and *z* (e.g. *benzin* (gasoline) instead of *bencin*) and in substituting *č* for *š* and *ž* (e.g. *Karnišnik* (proper noun) instead of *Karničnik*) fall into the category of Č/KONZ/menjava-drugo. All instances of *s-z* and *š-ž* are appropriately annotated with Č/KONZ/menjava-sz and Č/KONZ/menjava-šž.

## 1.2.9 Other consonant substitution = Č/KONZ/menjava-drugo

This category contains examples where the teacher corrected the incorrect use of a consonant within a word (a catch-all category encompassing all substitutions that occurred in the corpus' initial version fewer than 35 times). These problems often arise in transcribing (foreign) names.

### Č/KONZ/menjava-drugo:

- Ravno tako tudi Jošt Schwarzkobler | Schwarzkobler.
- Čez nekaj minut je stric oprtal svoj **nahrbtnit | nahrbtnik**.

Use of the letter *Đ* has proven especially problematic. When a student uses *đ* and the teacher changes it to *dj*, it is annotated as both a substitution and an omitted consonant.

### Č/KONZ/menjava-drugo in Č/KONZ/izpust:

- Bronja in **Đuro | Djuro** pa sta skupaj pobegnila.

Examples where students write a vowel instead of a consonant and vice versa are not treated as substitutions, but are annotated as dropped/superfluous consonants and dropped/superfluous vowels.

### Č/VOK/odveč in Č/KONZ/izpust:

- Djuro je Bronjo celo tako ljubil, da bi zanjo ubijal in **eradel | kradel**, vendar se moramo zavedati, da je tudi to zločin.

Examples featuring two consonants changing places are treated as metathesis.

### Č/SKLOP/premet:

- Torej je vsaj **dokačala | dočakala** lepo in mirno smrt tam med prijateljicami.

## 1.3 Labio-velar approximant w

The guidelines refer to “/w/ or the labio-velar approximant”, though from a linguistic standpoint it is imprecise or simplified terminology, as is the orthographic representation of such sound with <w>.

### 1.3.1 Word-initially = Č/W/začetek

These are examples where a teacher corrects the onset pronunciation to *u-* or *v-*. The annotated corpus contains examples of either *u-* used instead of *v-* (e.g. *ušeč*) or vice versa (e.g. *vsesti*). Combinations with other letter do not occur; if they had, we would group them in this category (e.g. *ga je \*ozela domov*).

### Č/W/začetek:

- **Vsedla | Usedla** sva se na tla in začel je govoriti.
- Zato bi raje pustila, da meni **uzamejo | vzamejo** življenje, kot da ga jaz jemljem drugim.
- A na žalost je mama silila ter me napadala z **uprašnji | vprašnji** in v tistem trenutku sem se raje zlagala.

Examples where the problematic phoneme is omitted are categorized as omitted consonants or vowels, but not also as an issue with word-onset *w*.

Č/KONZ/izpust:

- V nekaj sekundah pridrvi oče in **praša | vpraša** mamu kaj se dogaja.

### 1.3.2 Word-medial = Č/W/sredina

These examples show where teachers have corrected a labialized *w* within a word. Typically they feature the letter *v* instead of *u* and vice versa (*povdarjali*, *žuljenje*). Less often there also occur corrections changing *l-v* (*ponolno*, *sovze*) and *l-u* (*poune*).

Č/W/sredina:

- Tipična razsvetljenska ideja, saj so razsvetljenci **povdarjali | poudarjali**, da blaginjo v življenju dosežeš le z razumom **prevdarnostjo | preudarnostjo** in marljivostjo.
- To srce je **neuplivne | nevplivne** spravljal v obup, revščino, osamljenost, nekatere je to sovraštveno srce stalo celo življenja.
- Berk se kasneje kot komandir vojskuje na Hrvaškem, tik pred koncem vojne, pa se znajde v Beli krajini, kjer se **ponolno | ponovno** po dolgem letu sreča z Antonom.

This category also contains examples where a problem occurs within a proper name.

Č/W/sredina:

- Niti najmanj se ni oziral na grožnje in besede, ki mu jih je govoril **Zeus | Zevs**.

### 1.3.3 Word-final = Č/W/konec

This corrects spelling that ends in a labialized *w*. The annotated corpus primarily features problems with masculine participles, e.g. the use of *-l* instead of *-v* or *-u* (e.g. *je preklinjav*, *je plačau*). There are also problems with other words ending in *-v* (e.g. *zakonou*, *domol*) or *-l* (e.g. *dev*). If we found other similar examples (e.g. *-l* or *-v* instead of *-u*), they would also be placed in this category.

Č/W/konec:

- Polikarp je **preklinjav | preklinjal** v vseh možnih jezikih, med drugim tudi nemško, laško in španjolsko.
- Za to dejanje je Izidor drago **plačau | plačal**, saj ga je oče tako kruto kaznoval, da mu je odsekal prst.
- Po obisku pa Fatima ni odšla **domol | domov** vendar k sosednji hiši, kjer je bila zabava.

Examples featuring a reduction of the final vowel (e.g. in masculine participles) are annotated as problems with *w* as well as an omitted vowel. The same holds true for examples where the final vowel is additionally reduced, e.g. from *-u* to *-oi*.

Č/W/konec in Č/VOK/izpust:

- Vsak človek ima pravico pri Bogu do odpuščanja in če se je ta človek, ki je delal slabo, kesal in **spreobrnju | spreobrnil** se ga je Bog usmilil in mu odpustil ter ga sprejel k sebi.

- Zato se je župnik **odločo | odločil** da morata obe družini zvedet resnico, nakar je poklical Romea in Julijo kjer sta se pokazala.

Examples where the problematic phoneme is entirely omitted are categorized as omitted consonants, but not also as an issue with word-final w.

Č/KONZ/izpust:

- Se vidiva ka prideš **domo | domov**.

#### 1.3.4 Prepositional V = Č/W/v

This category includes all corrections of misused instances of the preposition v.<sup>1</sup> Primarily this appears as substitution for the letter u, though other examples (e.g. w) would also be grouped here.

Č/W/v:

- Ni pa čisto stoddstotno prepričan o izdajalstvu Katarine, saj **u | v** bistvu ni slišal kaj mu je ona govorila.

Examples where the incorrect usage of a preposition is also a function of a problem in writing words together/separately (e.g. *iz za* instead of *izza*, or *na mesto* instead of *namesto*) are appropriately annotated with both tags.

Č/W/v in Z/SN/SKUP-prisl:

- Gregor samega sebe **ubistvu | v bistvu** prekaša, saj ubiti osebo, ki ti veliko pomeni ni lahko.

### 1.4 Letter clusters

#### 1.4.1 A syllable is missing or is superfluous = Č/SKLOP/zlog

Here are grouped corrections where a letter cluster (usually a whole syllable) is either superfluous or is omitted. Common are examples where a string of letters is repeated in an ending. These examples are treated as a spelling issue and not a morphological one, as the ending was appropriately chosen, but a lapsus occurred in the doubling of the syllable.

Č/SKLOP/zlog:

- Vendar **menimin | menim**, da si zaradi ljubezni nebi nikoli stregla po življenju oziroma si ga vzela kakor so si ga Antigona, Hajmon in Evridika, čeprav nikoli neves, kaj si pripravljen storiti za ljubezen v določenih okoliščinah.
- Vsako leto se **zberejejo | zberejo** ljudje v Podpeci in gradijo gradove kralja Matjaža.
- Seveda, v obdobju miru bi na ista vprašanja, od istih oseb **dobili | dobili** popolnoma drugačne odgovore.

---

<sup>1</sup> Quantitatively this category does not feature more than 35 instances in the original corpus material; it is nonetheless maintained due to its specific nature and the fact that the examples herein do not belong to any other group.

If the correction makes it obvious that, in addition to an omitted or superfluous syllable, some other characteristic of the word is changed, this is also annotated appropriately. In the example listed the teacher added the missing "le" and removed the superfluous verbal prefix (žaželi – ža**že**leli), changing the choice of the verb in the process.

#### Č/SKLOP/zlog + B/GLAG/predpona

- Z mojo sestro sva si med jesenskimi počitnicami **žaželi | želeli**, da bi res odšle nekam na počitnice.

In terms of adding syllables, this was a typical example where 2 letters were missing in the middle or at the end of a word.

#### Č/SKLOP/zlog:

- Meursault je novico o smrti **sprej | sprejel** hladno.
- Ko se je popoldanski pouk končal je vedno **odhal | odhajal** iz šole med zadnjimi, ker se je bal, da ga bodo izivali, ker govori v dolenskem narečju.

Examples of missing syllables can often be easily separated from morphological corrections, since they do not involve typical problems located in endings. Whenever it is possible to interpret a problem morphologically as well, context and the nature of the other corrections aid in assessing which solution seems more likely. In the following example, for instance, we have annotated spelling and not the verbal ending, since there is a properly formed verb right next to the problematic word.

#### Č/SKLOP/zlog in Č/VOK/odveč:

- **Vzagaja | Vzgjajala** in učila jih je z vso svojo ljubeznijo; vendar mnenje od drugih ljudi ni nikoli spremenilo.

Examples that are unclear as to whether there was an incorrect choice in the use of a participle instead of a finite verbal form, or there was a spelling mistake in omitting a superfluous syllable, are annotated as morphological errors.

#### O/KAT/nedoločnik-osebna:

- Probleme ne smemo zavračati in se **dela | delati** kot, da se nič ni zgodilo.
- **Popredmeti | Popredmetiti** mora namreč sebe in sovražnika, se sprijazniti s tem, da je zločinec in izpolnjevati ukaze (ubijati).

### 1.4.2 Cluster LJ = Č/SKLOP/lj

These are examples where only *l* or *j* occurs where the cluster *lj* should be. The opposite also occurs, namely the presence of *lj* in places where only *l* is correct. These corrections are annotated doubly: once as a problem with the cluster *lj* and once as an omitted/added consonant.

#### Č/SKLOP/lj in Č/KONZ/izpust:

- Kajti sestra Ismena, bi s Antigonino odločnostjo in prepričanjem, da pokoplje brata, izgubila sestro, ki bi jo **kral | kralj** Kreon obsodil na smrt, torej je tudi Ismena postavljena pred isto dejstvo kot Antigona.



- Bori se za **bolši | boljši** položaj zamejskih slovencev, poučuje v slovenski šoli.

Č/SKLOP/lj in Č/KONZ/odveč:

- Veliki briljantni **valjček | valček** je podobna stvar, čeprav tukaj vidim bolj vzporednice s povojnim civilnim življenjem kot pa z razreševanjem starih zamer iz vojn.
- Do bližnjih ljudi moramo biti prijazni in nikakor se nam ni treba bati razočaranj, saj nam strah **zameglji | zamegli** pogledi in kmalu se lahko razvije sovrašтво.

Some examples (e.g. the above *valjček – valček*) can be interpreted as a vocabulary issue, but this interpretation is chosen only when it is clear from context that the word is actually being used in a different sense or meaning.

### 1.4.3 Cluster NJ = Č/SKLOP/nj

These are examples where only *n* or *j* occurs where the cluster *nj* should be. The opposite also occurs, namely the presence of *nj* in places where only *n* is correct. Metathesis sometimes also occurs, i.e. spelling *jn* instead of *nj*. These corrections are annotated doubly: once as a problem with the cluster *j* and once as an omitted/added consonant or as metathesis within a cluster.

Č/SKLOP/nj in Č/KONZ/izpust:

- Ismena se neizmerno veseli, da Polinejka ni, spodbuja **zvonenje | zvonjenje** zvonov, da oznanjajo mir.
- Na začetku romana je prikazan pogreb njegove žene Nadje, ki mu je predstavljala edini **življenski | življenjski**, intimni smisel v življenju.

Č/SKLOP/nj in Č/KONZ/odveč:

- Težave so bile tudi, ker se mati in Clairi, ter na začetku tudi Bubi niso hoteli naučiti **slovenjskega | slovenskega** jezika.
- Obema se poderejo cilji in ideali, le da Wertherju le **ljubezenjski | ljubezenski**, Črtomirju pa **ljubezenjski | ljubezenski** in verski.

Č/SKLOP/nj in Č/SKLOP/premet:

- V socialističnem času je sprva, takoj po vojni, zavel **vojn | vonj** svobode, vonj upanja.
- To ni nobenemu povedala razen Flori, katera ji je preko svojih prijateljev priskrbela zdravnika ki bi ji opravil ta poseg **zastojn | zastonj**.

Some examples (e.g. the above *vojn – vonj*) can also be interpreted as a vocabulary issue, but this interpretation is chosen only when it is clear from context that the student truly used the wrong word.

### 1.4.4 Cluster IJ = Č/SKLOP/ij

These are examples where only *i* or *j* occurs where the cluster *nj* should be. The opposite also occurs, namely the presence of *ij* in places where only *i* or *j* is correct.

These corrections are annotated doubly: once as a problem with the cluster *ij* and once as an omitted/added consonant or omitted/added vowel.

Č/SKLOP/ij in Č/KONZ/izpust:

- Na svetu je približno 6 **miliard | milijard** ljudi.
- Nima tudi nobenih posebnih bolezni, le nekaj težav s kožo v poletnih mesecih, zaradi **alergi | alergij** na trave.

Č/SKLOP/ij in Č/KONZ/odveč:

- Tudi danes lahko človek o svoji sreči odloča sam, vendar naše življenje vse preveč temelji na **materijalnih | materialnih** dobrinah.

There are also instances of *i* being used instead of *j* and vice versa.

Č/SKLOP/ij:

- V obeh zgodbah se zrcalijo okoliščine in problemi **socjalizma | socializma** pri nas.
- Simon Veber glavni lik v tej drami nič hugega sluteč odide na veselico kjer se napije, zjutraj pa se zbudi v **psihjatrični | psihiatrični** bolnišnici Svoboda osvobaja.

Problems with the insertion of *j* between a double *i* (e.g. *prijimek*) are annotated as shown below.

Č/SKLOP/podvojene in Č/KONZ/odveč in Č/SKLOP/ij:

- Je edini v družini \$\$\$ ki ima v **prijimku | priimku** še črko *j*.

If metathesis occurs within an expected *ij* cluster, the example is annotated with both appropriate annotations.

Č/SKLOP/ij in Č/SKLOP/premet:

- Roman je razdeljena na 14 **poglavji | poglavij** in Epilog.

At both the levels of both spelling and morphology we annotate examples where problems with the cluster *ij* are related to a transformation in the root or ending in the 1st and 2nd feminine declensions (SS 290-94), e.g. *ladja-ladij*, *Livia-Livij*, *bolezen-boleznijo*, etc.

#### 1.4.5 Doubled letters = Č/SKLOP/podvojene

This group encompasses all corrections arising due to a doubled letter. These can be words with a doubled letter that the student failed to write, or examples where the student did write a doubled letter where only one belongs. We also annotate these examples to show whether the mistake was an omitted vowel/consonant, its addition, or even its substitution.

Č/SKLOP/podvojene in Č/KONZ/izpust:

- V večerni **odaji | oddaji** je bila Davidova smrt poročana.
- Laert je bil bolj odločen v maščevanju, in Hamleta iz same jeze hoče ubiti, zato ga **izove | izzove** na dvoboj.

Č/SKLOP/podvojene in Č/VOK/izpust:

- Vendar župnik na koncu Jermanu **vseno | vseeno** prizna, da je ravnal prav.

Č/SKLOP/podvojene in Č/KONZ/odveč:

- **Oddlomek | Odlomek** govori o odločnosti nesrečnega Danila kako se je zatekel v sanitarijo da bi si nekako povrnil novih moči.
- **Narodna | Narodna** univerzitetna knjižnica

Č/SKLOP/podvojene in Č/VOK/menjava-uo:

- Ampak preden rečemo nebom se **suočil | soočil** z njim ga moramo dobro analizirat.

Č/SKLOP/podvojene in Č/KONZ/menjava-drugo:

- Ne deloval je instinktivno, nagonsko kot **razzarjena | razjarjena** žival.

Other examples that go in the category of doubled letters are ones where *j* is inserted between doubled vowels.

Č/SKLOP/podvojene in Č/KONZ/odveč:

- Dan je bil naporen a se je **vsejeno | vseeno** splačalo.

#### 1.4.6 Metathesis = Č/SKLOP/premet

Here belong those corrections where all the right letters occur within a word but where 2 of them have switched places. Usually either two vowels or two consonants switch places.<sup>2</sup> If the correction in question occurs within one of the described problematic clusters (*lj, nj, ij*), it is additionally annotated.

Č/SKLOP/premet:

- Sprva se začne udejestvovat v slikarstvu, potem to **uposti | opusti** in najde smisel v literaturi.
- Torej je vsaj **dokačala | dočakala** lepo in mirno smrt tam med prijateljicami.

Č/SKLOP/NJ + Č/SKLOP/premet/

- Nato pove še svoji prijateljici Flori ona pa ji prek vezi poišče zdravnika, ki bo naredil splav **zastoin | zastonj** saj Špela ni imela denarja.

### 1.5 Variable (allophonic) prepositions

#### 1.5.1 Preposition s/z = Č/PREDL/sz

These mistakes are in the incorrect choice of allophonic variants of the preposition *s* and *z* (this category could also be grouped at a higher level, but by analogy it is currently listed as a spelling mistake).

Č/PRED/sz:

- Po eni strani se smili samemu sebi, saj vse opravičuje **z | s** smrtjo svoje žene, ki mu je vse pomenila.
- Tragedija je igra, ki se konča **z | s** smrtjo glavne osebe oz. konec je nesrečen.

<sup>2</sup> This category does not have a significant amount of occurrences in the corpus, but it was nonetheless kept, as literacy experts maintain that it shows the result of very specific deficiencies, e.g. of dyslexia.

- Sicer pa niti znanstveniki ne morejo dokazati, kdo je imel prav, saj je bila odločitev **s | z** Antigone in Kreonove strani hkrati pravilna, vzorna za druge in etično dobra.

### 1.5.2 Preposition k/h = Č/PREDL/kh

These mistakes are in the incorrect choice of allophonic variants of the preposition *k* and *h* (this category could also be grouped at a higher level, but by analogy it is currently listed as a spelling mistake).

Č/PRED/kh:

- Bolj ko gre zgodba **k | h** koncu, bolj sta Romeo in Julija zaljubljena.
- Stražarji jo zjutraj odpeljejo **k | h** Kreonu, ki jo zaslišuje, vendar je Antigona še zmeraj trdno prepričana v pravičnost svojega dejanja.
- Odneseš me je **h | k** socialni delavki, kamor me je položil na zločljivo posteljo, kjer sem se zbudil.

## 2 MORPHOLOGY

The morphological level includes corrections to inflection and, partially, word formation. The distinction between morphology and spelling or vocabulary is not always a trivial one. A few more complicated examples are discussed in Chapter 8. An additional challenge awaits in corrections that encompass more than one word in a phrase (e.g. a correction to a case in a noun phrase with an agreeing adjectival qualifier), as these are segmented in the current corpus in an inconsistent manner: often, but not always, they are annotated as 2 or more separate corrections.

- Menim da je v knjigi zelo vidna krivica, ki sta jo bila deležna Romeo in Julija, ki sta za zgladitev sporov plačala z **svojim | svojima življenjem | življenjema**.

These examples are annotated for category on both of the corrected words.

### 2.1 Categorical corrections

It is characteristic for this group that the corrected word is otherwise correctly formed (occurs within a declension or conjugation table), but it is categorically incorrect for some reason – e.g. shows the wrong case, number, person, etc.

#### 2.1.1 Case: genitive-accusative = O/KAT/sklon-rt

These are very common mistakes where the teacher corrects the genitive case into accusative or vice versa.

O/KAT/sklon-rt:

- Vem, da se usode ne da predvideti, **jo | je** spremeniti, a vendar naj bi se vsak za svojo srečo potrudil.
- Kaplan doseže svoj namen saj vidi, da **vsakega | vsak** trenutka vse več src pritrjuje na cerkvena vrata in vidi da se ljudem od radosti in narodne zavesti lesketajo oči.

Here we also group problems of declension, such as those that occur with the negative pronoun *nič*. This can be understood as a matter of an indeclinable word, but here it is interpreted as an issue with the genitive and accusative.

O/KAT/sklon-rt:

- Sam menim, da je prav cerkev kriva za številne vojne, lakote in nezadovoljstvo ljudi samo kaj ko sem le eden od peščice, ki meni tako in ne morem **nič | ničesar** spremeniti.

Since the examples here are few and far between, currently they are also grouped as mistakes in determining animacy, as indicated in the example below.

O/KAT/sklon-rt:

- To velja za Silvana Kandorja, **glavnega | glavni lika | lik** romana Senčni ples.

### 2.1.2 Case: dative-locative = O/KAT/sklon-dm

In these examples teachers have corrected the dative to locative or vice versa.

O/KAT/sklon-dm:

- To pa ga pripelje no razklanosti v **samemu | samem** sebi.
- Odlomek je epski, saj v njem pisatelj pripoveduje o **temu | tem**, kako je bilo, ko se je vrnil ata domov.
- Ko se je pozdravil, se zahvalil **zevsovem | zevsovemu** sinu za pomoč.

Sometimes it is unclear whether the original (corrected) form was in the dative or the instrumental. In annotating such mistakes we assume the type of substitution that is most typical.

O/KAT/sklon-dm:

- V priliki o izgubljenem sinu imamo prav tako motiv dveh bratov, pri **katerima | katerih** se ponovi njun medsebojni odnos.

How typical a given linguistic issue is also serves as a guide in similar ambiguities. In the example below, e.g., we assume that the correction is at the level of case and not vocabulary (i.e. that the statement was not originally meant to be something like: *Jazon je odporen proti vsem ljudem*).

O/KAT/sklon-dm:

- Jazon je bil odporen proti **vsem | vsemu**.

### 2.1.3 Case: locative-instrumental = O/KAT/sklon-mo

In these examples teachers have corrected the locative to instrumental or vice versa.

O/KAT/sklon-mo:

- Jaz osebno ne bi nikoli se tako gnal za **nečem | nečim**, če mi ne uspe mi pač ne uspe — bom pa drugič.
- Vse skupaj se začne z **moškem | moškim**, ki se je na avtobusu peljal proti domu.

## 2.1.4 Other case substitutions = O/KAT/sklon-drugo

Here belong examples where the teacher corrects the morphology of the chosen case.

O/KAT/sklon-drugo:

- Jaz sem se odločil za trening vsak dan, čeprav mi to vzame veliko časa in energije sem vesel, da sem se tako odločil, saj sem prišel tako, do veliko uspehov in sem zelo blizu **najboljšim | najboljših** v Sloveniji.
- Roman v **dvaintrideset | dvaintridesetih** poglavjih Senčni ples, pa je napisal Alojz Rebula.
- Odlomek **Nezakonsko | Nezakonska** mati je napisal France Prešeren in spada v obdobje Slovenske romantike, ki traja od 1830 do 1849.

The word *več* can have (including in terms of its prescribed morphosyntactical attributes) the role of adverb, particle, or quantifier. Whenever there is an example where the word *več* is corrected into *večih* or vice versa, we interpret it as a quantifier and annotate it as a correction to case.

O/KAT/sklon-drugo:

- Naša družba je prvič v **večih | več** letih dobila „vodjo“, nekoga, ki je iztopal, ki je bil nekaj več kot vsi ostali, nekoga, ki je imel glavno besedo.
- V tebi se razraščam je združila poezijo **več | večih** avtorjev, nastalo v skoraj dvesto letih.

Substitutions in the case forms of the words *mati*, *hči*, and *otrok* are additionally annotated as morphologically irregular words. Problems with case include substitutions of forms that actually exist, while non-existent forms (e.g. *matero*) are annotated as paradigm problems.

O/KAT/sklon-drugo + O/DOD/besede-mati-hči:

- Prva **mater | mati** je tudi pravična, saj je vedela resnico, jo govorila.
- Ker če bi nehal raziskovat očetovo smrt, ga ljudje ne bi imeli za blaznega in Polonij bi pustil **hči | hčer** in Hamleta pri miru.

O/KAT/sklon-drugo + O/DOD/besede-otrok:

- To, dobro in slabo vidimo tudi v osnovnih šolah med **otroci | otroki**.

Currently under case we also list corrections in the use of the 2nd masculine declension. When the correction addresses a morphological variant (both forms are used in standard language), it is additionally annotated for variability.

O/KAT/sklon-drugo + O/DOD/variante:

- Ni se bal **Volodja | Volodje**, tako kot so se ga drugi.

Whenever a case correction occurs as a result of one of the teacher's other corrections (e.g. corrections at the level of vocabulary, as in the example below), it is annotated as a related correction (more on this in Chapter 6).

P/OBL/drugo:

- Ko Kreon le spredvidi kaj je storil je bilo že prepozno škoda je že bila storjena in kreon je v bolečinah in **v | z občutku | občutkom** krivde moral živeti naprej.

### 2.1.5 Number: singular-plural = O/KAT/število-em

These examples show where teachers have corrected a singular form into a plural form and vice versa. This tag joins all examples of such a substitution, regardless of part of speech.

#### O/KAT/število-em:

- A hlapci so se samo skrili za **grmovja | grmovje**, kjer jih je Martin takoj opazil.
- **Rezultati | Rezultat so | je** bili 2 proti 0 za mene in XImeX.
- Tako hitro se je vse spremenilo, ampak to **opazimo | opazim** šele, ko se pogledam v sliko ali pa mi to kdo drug pove.

Also included here are corrections to the number of the possessor in possessive pronouns.

#### O/KAT/število-em:

- Po mojem mnenju je ta ljubezen res čutna in prava in je res škoda, da ni **njun | njen** konec drugačen.

### 2.1.6 Number: dual-plural = O/KAT/število-dm

These examples show where teachers have corrected a plural form into a dual form and vice versa. This tag joins all examples of such a substitution, regardless of part of speech.

#### O/KAT/število-dm:

- Opazili smo **jih | ju** v bližini Ankarana, Kopra, Pirana, ter Savudrije.
- Enkrat se **moramo | morava** dobit in pogovarjati o naših starih dogodivščin.
- "No ne buljita preveč, saj veva, da sva **najlepše | najlepši**," je prevzetno rekla Rozamunda.

Issues in composing verbal forms are considered a matter of number if the problem is only at the level of individual constituent, e.g. in the following e.g. the auxiliary is right but the participle was corrected:

#### O/KAT/število-dm:

- S prijateljico sva se **odločile | odločili**, da se ne bova **udeležile | udeležili** pouka pri učiteljici, ki nama je presedala vsa leta šolskega življenja.

In examples like the following it was not the number that was corrected, but the non-standard ending; these examples are group into the appropriate subcategory of paradigmatic mistakes.

#### O/PAR/glagolska-končnica:

- Mislili **sve | sva** da **bove | bova** zelo lahko prišle do Amerike in na hitro obogatele, ampak videle **sve | sva** da to ni lahko, ampak sva vstrajali.

When it is clear from the correction itself that a question of number or other category is the issue, i.e. gender, a decision is made based on the context.

O/KAT/število-dm:

- Olimpijske igre se izvajajo vsaka štiri leta. Na **vsaka | vsaki** dve leti si tako lahko ogledamo ene.

There are also examples where in addition to annotating number another annotation is also made. The following example shows the incorrect use of number and a poorly formed ending:

O/KAT/število-dm + O/PAR/neglagolska-končnica

- Ko sta se poti ločili proti **njihovimu | njunima** domovoma, sta si morala razdeliti dobiček.

### 2.1.7 Number: single-dual = O/KAT/število-ed

These examples show where teachers have corrected a single form into a dual form and vice versa. This tag joins all examples of such a substitution, regardless of part of speech.

O/KAT/število-ed:

- Meščanstvo in ožji del družine se z njegovo odločitvijo **ni | nista strinjalo | strinjala**.
- Tudi njegovo osvoboditev **onemogoča | onemogočata** oblast in država.
- Oblekli sva najbolj **razkošno | razkošni obleko | obleki** kar sva jih imeli.

### 2.1.8 Gender = O/KAT/spol

These examples show where teachers have corrected a form due to the chosen gender.

O/KAT/spol:

- Špela in Flora sta se **odpravila | odpravili** na prvi šolski ples.
- Govori o dekletu, ki je **zanosila | zanosilo** in fant jo je zapustil, zato je bila nezakonska mati.

The category of gender also contains problems where a possessive adjective is formed from a personal name according to the paradigm for the wrong gender. This does not include examples where the choice of noun was corrected (*junak-junakinja*); these are considered vocabulary errors.

O/KAT/spol:

- **Antigonovo | Antigonino** dejanje se Kreonu ni zdelo prav nič lepo.

Examples of the sort *nadaljnjo življenje* are also grouped here, as they are usually bound exclusively to the singular of the neuter gender in the nominative or accusative case. This is not a typical example of a gender-related problem, but it is possible to superficially describe it as such (*zahvaljujem se za vašo pomoč* vs. *zahvaljujem se za vašo dovoljenje*).

O/KAT/spol:

- Zaupal je vanjo in v Boga, ki mu je dajal moč za **nadaljnjo | nadaljnje** življenje.



- Zahvaljujem se vam za **vašo | vaše** dovoljenje.
- Truplo je bilo **vso | vse** razžrto od psov.

### 2.1.9 Aspect = O/KAT/vid

This category is for corrections to verbal aspect, i.e. to the imperfective or perfective form of a verb.

O/KAT/vid:

- V odlomku iz Hamleta se je namreč **osredotočal | osredotočil** na njega samega.
- Ker bi radi vse čimhitreje natipkali, skrajšujemo besede ali pa **prevzamemo | prevzemamo** iz drugega jezika.
- Ko Romeo in Julija povesta celotno zgodbo in da sta zaročena, se družini pobotata in **podpirata | podpreta** svoja otroka Romea in Julijo.

Examples where the change in aspect is related to the verbal prefix are currently considered as vocabulary mistakes as well, and are appropriately ascribed both annotations.

B/GLAG/predpona + O/KAT/vid:

- Ko sem se **tuširal | stuširal** sem šel v dnevno sobo.
- Drugače pa je bil družinski človek in vedno **pomislil | mislil** na druge potem pa na sebe.

### 2.1.10 Tense = O/KAT/čas

In these examples teachers have corrected the verbal tense. For example, it is common to have corrections from the past tense to the present.

O/KAT/čas:

- Tudi jaz sem ji povedal, kako zelo **sem jo ljubil | jo ljubim** in tako sva se pogovarjala §§§
- Žal se je ljubezen tragično končala saj **Meta umre | je Meta umrla** v naročju svojega ljubljenega.
- Odločil sem se, da jih **bom pobil | pobijem** ujetnika pa rešil.

Here are also grouped examples where a vocabulary choice was corrected, but where it is undoubtedly clear that the teacher also corrected the verbal tense.

O/KAT/čas + B/GLAG/drugo:

- V njem pa **je napisal | sporoča**, češ da se Baronica z nekom sestaja, da bi tako zbudili njegovo ljubosumje.

This category also contains examples with an omitted participle, which is expressed as a substitution of verbal tense.

O/KAT/čas:

- To sta **bili | \_\_** osebi, ki sta hoteli osvojiti Nežko.
- Kreon je **bil | \_\_** zato pokazan kot človek, ki mu je oblast **bila | \_\_** bolj pomembna, kot pa sorodstvo.

### 2.1.11 Person = O/KAT/oseba

This tag combines examples where the teacher corrected the verb's person.

#### O/KAT/oseba:

- Mislim torej, da so nezakonske matere ženske, katerim se ljudje lahko le klanjamo in se **zgledujejo | zgledujemo** po njih.
- Mislim, da če človek ne goji ljubezni, iskrenosti in zaupanja do drugih, tudi svobode ne more čutiti, saj potem **nosiš | nosi** vedno neko tesnobo in težo v svojem srcu.

If in addition to correcting person there is also another evident correction (e.g. tense or number), other appropriate tags are also ascribed.

#### O/KAT/oseba + O/KAT/število-em:

- Če v življenje ne **vključimo | vključiš** malo tekmovalnosti, sovraštva, si vedno na zadnjem mestu.

### 2.1.12 Reduced infinitive = O/KAT/nedoločnik-kratki

In these examples teachers have corrected the reduced infinitive into its full form. This category must be distinguished from the category dealing with the substitution of infinitive and supine. Corrections to a reduced infinitive occur in sentences without a verb of motion and deal with the dropping of a verb-final *-i*.<sup>3</sup>

#### O/KAT/nedoločnik-kratki:

- Medtem Nežka pove baronici, da jo baron osvaja in ta želi **ukrepat | ukrepati**.
- Želi se mu **odkupit | odkupiti**, ker ga za časa življenja ni imela dovolj rada.

### 2.1.13 Infinitive and supine = O/KAT/nedoločnik-namenilnik

This tag joins corrections of the infinitive into the supine.

#### O/KAT/nedoločnik-namenilnik:

- Za pljučnico sem zbolel, ker sem šel s prijatelji na stadion v Vipavi **igrati | igrat** nogomet, ko se je usul močan dež.
- Kasneje se je šel Hamlet **pogovoriti | pogovorit** z materjo, vendar je v njeni sobi bil Polonij, ki se je skril za zaveso in ga je Hamlet zabodel, ker je mislil, da je on Klavdij.

---

<sup>3</sup>This category of reduced infinitive is kept separate due to its formal similarity with problems related to the supine, although due to the non-standard nature of the corrected form it might be better to group it among paradigmatic correction, or even, it could be argued, at the level of spelling.

### 2.1.14 Infinitive and finite verb = O/KAT/nedoločnik-osebna

Here we group examples where the teacher corrects the use of the infinitive into a finite verbal form. These examples often show syntactical contamination, where the use of certain verbal forms in a sentence beget further mistakes.<sup>4</sup>

#### O/KAT/nedoločnik-osebna:

- On kot veliki avanturist, ki ni bil sposoben pokazati čustev, intelektualec se je znašel v zmedi vojne in nekako pod vplivom razmer tudi **delovati | deloval**.
- Takoj, ko vstopis vstaneta Sonja in Jaka ter mi začneta **ploskata | ploskati**.
- **Popredmeti | Popredmetiti** mora namreč sebe in sovražnika, se sprijazniti s tem, da je zločinec in izpolnjevati ukaze (ubijati).

Corrections at the level of infinitive–finite verbal form are often the result of other syntactical corrections, as shown by the example below. These examples are grouped among related corrections (more on those in Chapter 6).

#### P/OBL/drugo:

- Pa vendar mi je bil ta lik zanimiv, saj ga je Linhart najbrž hotel takega ustvariti, **saj je hotel s to | da bi s** komiko nravi še bolj **poudariti | poudaril**, da je to njegovo delo komedija.

### 2.1.15 Reflexivity = O/KAT/povratnost

This group covers examples where the teacher corrected *si* into *se* or vice versa. This problem could potentially be considered one of proper case, but do to its frequency and specific nature we have kept it in a separate category.

#### O/KAT/povratnost:

- Ofelije **si | se** je začel izogibati in jo ignorirati.
- Spoznala sta se na potovanju, potem sta **se | si** dopisovala prek interneta in za počitnice je ona šla k njemu.

Here are also grouped examples where the reflexive pronoun *se* or *si* is omitted or is written superfluously. Special attention must be paid here, as similar omission or superfluous additions are typically grouped at the syntactical level.

#### O/KAT/povratnost:

- Ko sem ga prehitel pa **se | \_\_** mi je spodrsnilo in na vrat na nos sem padel s kolenom na betonski rob in zakričal.
- Tako **si | \_\_** je v naši tragediji sebično vzela k sebi Hajmona, ki je imel še celo življenje pred seboj.

---

<sup>4</sup>This category is currently grouped at the morphological level, though it partially overlaps with spelling. This occurs fewer than the threshold 35 times in the corpus, but it is kept separate due to its specific nature.

- Nato **sem se | sem** vstala in veselo odšla proti vrhu.

#### 2.1.16 Mood = O/KAT/naklon

This category encompasses all problems related to the verbal mood: indicative, conditional, and imperative. The majority of these mistakes are related to the introduction of the conditional. Longer structures are often involved in such corrections, wherefore this problem approaches the realm of syntax. When deciding how to handle a certain correction (which parts go together and which parts should be treated as a separate correction), we follow the codes <err> and <corr>, even if the current segmentation is not necessarily optimal. The word *naj* is often involved in the erroneous string and it does not get specially annotated as an added/inserted word.

##### O/KAT/naklon:

- Zato si je Odisej izmislil pameten in prebrisan načrt, da **bo | bi** oslepil Polifema.
- Imam pa občutek da bubi bolj, kot pa pripisovanju krivde, posveča svoje misli k temu, kako **lahko on izboljša | bi lahko izboljšal** svoje življenje in življenje njegove družine v okoliščinah v katerih so.
- Šlo je za veliko vsoto denarja, ki **so si | naj bi si** ga razdelili 3je možje.

Questions of tense are not annotated specially when the issue revolves around the use of a participle in the conditional mood (e.g. *bi umrla* vs. *umre*). If another change is also present, e.g. at the level of aspect, gender, etc., that change is annotated as well.

##### O/KAT/naklon in O/KAT/vid:

- Če ne najdem kranjske, bom moral žvižgat, zakaj nemško ne znam prav," je Matiček izrekel in se boril, da **se slovenščina uporablja | bi se slovenščina uporabila** tudi v uradih.

Further, in this category are grouped (the otherwise rare) examples where the teacher in changing the verbal form also changes an affirmative statement into a negative one or vice versa (the affirmative, negative, and interrogative are covered in grammar by the indicative mood).

##### O/KAT/naklon:

- V glavi so ga nenehno ovirali glasovi umirajočega Jošta :„Pridem pote !” a se za njih ni zmenil, dokler **je | ni** na njegovo domačijo prišla gospa po imenu Paseverica.

#### 2.1.17 Voice = O/KAT/način

These examples show where teachers have corrected passive voice into active voice and vice versa. Like the previous category, this one transcends the level of word (perhaps it would be best to treat it at the syntactical level, but for now it remains a morphological question). Corrections in this category are usually rare, but nonetheless complex or multifaceted and often require additional annotation.

##### O/KAT/način + B/MEN/polnopomenska-v-zaimek:

- **Delo je bilo igrano | Igrali so ga** na Dionizijevih igranj in zato jo imenujemo tragedija.

#### 2.1.18 Pronominal form = O/KAT/oblika-zaimka

This category groups all examples where the correction targets the choice between the clitic, prepositional, and stressed forms of the personal pronoun.

##### O/KAT/oblika-zaimka:

- Nežka bi morala Matičku zaupati, da se bo gospa (preoblečena **v njo | vanjo** sama srečala z baronom.
- Penelopa in Odisej sta si imela toliko za povedati, da **njima | jima** je Atena podaljšala noč.

#### 2.1.19 Definiteness = O/KAT/določnost

These corrections change the form of a definite adjective or pronoun into an indefinite one or vice versa. Although in certain examples these corrections could be grouped at the level of spelling (as the omission/addition of a word-final vowel, e.g. *bogati kmet*), when in doubt we decide on the morphological level.

##### O/KAT/določnost:

- **Nov | Novi** kralj je postal mož plemenite, poštene krvi, ki je tudi državi dejal, da če bi Hamlet preživel, bi bil on prava oseba za kraljevski stolček.
- Meta rodi še pet otrok, Ožbej pa je **vsaki | vsak** dan bolj pijan.

This category also contains corrections to the pronoun *neki – nek*.<sup>5</sup>

##### O/KAT/določnost:

- Na **nek | neki** način sva le dosegli svoje sanje.

#### 2.1.20 Comparison = O/KAT/stopnjevanje

Under problems with comparison are grouped examples where the teacher corrected a synthetic form of comparison into an analytical form or vice versa. This includes examples where the teacher corrected a comparative form into a non-comparative form or vice versa. The addition or omission of the adverb *bolj* or *najbolj* here is not treated as a separate correction.

##### O/KAT/stopnjevanje:

- Veliko bolj **prepričlivejša | prepričljiva** so seveda dejanja Črtomira, ki dela dobro in je zame zares pravi junak, glede ljubezenske vdanosti.

---

<sup>5</sup> Etymologically speaking this is a matter of reduced word-final *-i*, wherefore they could easily be considered a spelling issue. In the contemporary form of the language, however, it is perceived as a question of (in)definiteness, and formally speaking the correction is similar to other corrections to (in)definiteness, so it is sensible to keep it at the morphological level.

- Čedermacev govor je bil **najpreprostejši | preprost**, da so ga lahko razumeli, vendar se je trudil govoriti v lepem jeziku.
- Od tistega trenutka naprej, ko je moril je postajal le še **bolj slabši | slabši** in krutejši.

If a spelling mistake also occurs in the example, it is only annotated in the correction evidently relates to this as well. In the first example below, for example, only the correction to comparison is annotated, while spelling is annotated in the second one.

O/KAT/stopnjevanje:

- Tokrat bo **bol objektivni | objektivnejši** in bom moja moja čustva prihranila za nadaljevanje.

O/KAT/stopnjevanje + Č/SKLOP/lj + Č/KONZ/odveč:

- Tam so imeli dosti **bolj lepo | lepše** življenje.

We also group in this category those problems where the teacher corrected the comparative or superlative form of the adverb *raje*, which is problematic due to the morphological variability in usage. Since *raje/rajši* and *najraje/najrajši* are all legitimate morphological variants, they are accordingly annotated as variants in the corrections. The forms *rajše* and *najrajše*, however, are annotated as errors in comparison and as paradigmatic problems of non-verbal endings.

O/KAT/stopnjevanje + O/DOD/variante:

- Od hrane ima **najrajši | najraje** pico, makarone in palačinke, od pijače pa Coca colo, Fruc in Sprite.
- Vnekatih primirih je **bolje | boljše** prilagajanje, v drugih da se držiš svojih načel.
- Včasih se ljubezen prehitro konča čeprav bi lahko trajala **dalj | dlje**.

O/KAT/stopnjevanje + O/PAR/neglagolska-končnica:

- Zato sem jast zelo osamljen in sem **najrajše | najraje** \$\$\$sem\$\$\$.

## 2.2 Paradigmatic Corrections

Paradigmatic corrections join all examples where forms occur in the text that are not part of any standard paradigms. Often at work here is word formation based on false analogies or under the influence of regional or dialectal morphology.

### 2.2.1 Verbal root = O/PAR/glagolska-osnova

This category encompasses examples where the teacher has corrected a verbal root that was formed by false analogy, e.g. according to the infinitive form or by using inappropriate morphemes for (im)perfectiveness.

O/PAR/glagolska-osnova:

- Lojzkin oče se je **dobijal | dobival**, mama njena pa je svoje ljubimce kar domov pripeljala.
- Čudilo ga je tudi, ker je Albin **rastel | rasel** veliko hitreje kot on in da je bila tam vojna.

- Moja babica je preprosta oseba, ki ji ni težko kdaj **poprimiti | poprijeti** za delo in pomagati.

### 2.2.2 Verbal ending = O/PAR/glagolska-končnica

This category encompasses examples where the teacher has corrected a verbal ending formed by false analogy.

O/PAR/glagolska-končnica:

- Že od zmeraj, hočem biti vsaj malo bolj glaven kot drugi, vendar to me **pelja | pelje** samo v težave in konflikte.
- Antigona se je kljub temu odločila, da Polinejka **pokopa | pokoplje**.

A part of these errors is connected to the paradigms of athematic verbs, i.e. to the endings *-(s)ta* or *-(s)te*, which refer to the third person in the dual and plural.

O/PAR/glagolska-končnica:

- Adam povabi Fatimo v svojo sobo in **imasta | imata** prve spolne odnose.
- Sklenila sta, da se **bota | bosta** poročila.

Among those mistakes most related to regional influence, the occurrence of the ending *-ma* instead of *-va* in the first person dual, and the usage of the ending *-ve* instead of *-va* for the first person feminine dual.

O/PAR/glagolska-končnica:

- Komaj čakam, da te spet vidim, da odhitiva malo naokrog, na sprehode, v kino in skupaj **opravima | opraviva** v domačih opravilih.
- **Šle sve | Šli sva** na sprehod in povedala mi je vse, kar se je dogajalo v družini.
- Spominjam se najinega pobega, ko **sve | sva** se odločili, da **bove | bova** postali slavni.

### 2.2.3 Non-verbal root = O/PAR/neglagolska-osnova

This category gets all the corrections made to erroneous forms where the problem occurs in the root of the non-verbal word (i.e. noun, adjective, adverb, quantifier, etc.). It is typical of this category that the student failed to extend the root (e.g. with *-j-*, *-t-*, *-n-*, *-ov-*), or extended roots that should not have been extended (e.g. with *-ot-*, *-at-*). For more information on root transformations c.f. SS 280-81.

O/PAR/neglagolska-osnova:

- To sem povedala **Jakatu | Jaku** in Anžetu a takrat mi je Anže začel govoriti lepe besede.
- Tako tudi želi Kreonu dokazati, da ne more zakone ljudem vsiliti in da zakone določajo **bogi | bogovi**.
- Kaos se konča z energičnim nastopom **Kalanderja | Kalandra**, ki po svojih najboljših močeh brani Jermana.
- Izidor je mirna oseba, zadržane narave in je dobrosrčen, saj to tudi opazimo v odlomku, kjer oče Izidorju odseka pol mezinca, a **Izidorov | Izidorjev** odziv ni takšen, kot bi ga pričakovali.

Also grouped here are examples where the problem arose due to significant irregularities in the declension paradigm, e.g. *človek* and *uho*, where the nominal root itself transforms (SS 271).

O/PAR/neglagolska-osnova:

- Lojzka je bila primorana nemočno opazovati nečedna dejanja svoje matere in očeta, skratka **človekov | ljudi**, ki sta ji podarila življenje in imela dolžnosti pravilne vzgoje, biti lep zgled in postaviti otroku temelje prihodnosti.
- Počasi sem spravila burek v vrečko, z roko vzela slušalko iz **uha | ušesa** in sledila glasu.

Examples where the teacher corrects a form listed in reference resources as a variant is additionally annotated as such (cf. Chapter 2.3).

O/PAR/neglagolska-osnova + O/DOD/variante:

- Mama je cele dneve z hčerko pri zvezkih in ne skrbi niti malo za hiho ali pa za svoja dva **sina | sinova**.

Here we also group the less common examples where the transformation in the root is not registered in the reference grammar due to the non-standard nature of the form.

O/PAR/neglagolska-osnova:

- Odlično sva se razumeli in tudi odnos njenih staršev je bil do mene, do moje prijateljice in navsezadnje tudi drug do **druzga | drugega** tako ljubeč, poln topline.

## 2.2.4 Non-verbal ending = O/PAR/neglagolska-končnica

This category contains examples where the teacher has corrected the use of a non-standard ending in a part of speech that is not a verb. Often here there occurs the problem of failing to heed a transformation in the ending or of simply forming the ending by false analogy.

O/PAR/neglagolska-končnica:

- Njena ljubezen z Ožbejem je bila tako močna, da ju vse **kazne | kazni**, ki jih je dobila Meta, niso mogli ločiti.
- Od takrat naprej sta družini živeli v miru in Romeo ter Julija z **ljubezno | ljubeznijo** do konca dni.
- Lahko je reči, ampak ko prideš v položaj in izbiraš med dvema **možnostima | možnostma** je drugače.
- Zaljubila sem se **vanjga | vanj**.

Also grouped in this category are endings that are typically recognized in colloquial speech as regional varieties of Slovenian. If these endings exist in the standard paradigms, too, then extra attention is needed. If a correction cannot be placed at the categorical level and if the metadata confirm that it could concern an instance of regional usage, such an example is grouped here.

O/PAR/neglagolska-končnica:

- Koje bil star 15 let, in po **končanoj | končani osnovnoj | osnovni** šoli, se je odločil, da bi se učil za gradbenika.



- Se ne upaš učiti tolko dolgo kot si bil po **računalniki | računalniku**, naj ti on pomaga napisati domačo nalogo!
- Anne je umrla okoli 2 meseca pred koncem vojne, njen fant Peter pa 3 ali 4 **dana | dni** pred koncem vojne.

Similarly, here belong problems with the words *mati*, *hči*, and *otrok*, where students form words according to imaginary or non-existent paradigms (the forms *materjo* and *hčerjo* are considered ending transformation and not root extension, SS 291). Regular substitutions (e.g. *mati-mater*, *hči-hčer*, *z otroci-z otroki*) are treated as categorical problems. Substitutions in the forms of the difficult words *mati*, *hči*, and *otrok* are additionally annotated as problematic words.

O/PAR/neglagolska-končnica + O/DOD/besede-mati-hči:

- Kolonija predvsem skrbi, da Hamlet kaj ve o njegovem zločinu in zelo izkorišča svojo ženo, zato se odnos med **matero | materjo** in sinom še bolj ohladi.
- V Idriji je spoznal svojo ženo XImeX s katero je imel dva sinova in eno **hčero | hčerko / hčer**.

O/PAR/neglagolska-končnica + O/DOD/besede-otrok:

- Ko je pred cerkvijo doživel prvo ponižanje, sem videla, kako zelo podobni smo si z **otrocmi | otroki** iz tistega časa.

### 2.2.5 Epenthetic vowel = O/PAR/neobstojni-vokal

This category contains corrections to the transcription of the epenthetic vowel or the omission of an epenthetic vowel in the word form.<sup>6</sup> These are primarily examples related to the semivowel (also referred to as “schwa”) *e* (/ə/). When in doubt as to whether a given example should be ascribed to the level of spelling, it serves to compare the word’s root form vs. its declined form (*Abel* – *Abela* or *Abla*).

O/PAR/neobstojni-vokal:

- Zakaj smo nekateri **bolani | bolni** in drugi zdravi?
- Bog se je zmenil samo za **Ablovo | Abelovo** daritev, zato je bil Kajen nekoliko ljubosumen.
- Zevs je vedel, da ga hoče pretentati in je zato izbral slabši kup s **kostemi | kostmi**.

### 2.2.6 Umlaut and cč = O/PAR/preglas-in-cč

This category features corrections where a student either ignored or hypercorrected umlaut. The problem is common among possessive adjectives formed from proper nouns and with nouns in the instrumental.

O/PAR/preglas-in-cč:

- V tej zgodbi se mi zdi **Prometejov | Prometejev** upor pozitiven.

<sup>6</sup>Numerically speaking, this category does not meet the threshold 35 examples, but it is nonetheless preserved for practical purposes.

- Nekega dne smo se pogovarjali s **stricom | stricem**.

In forming possessive adjectives there are also problems with transformations of *c* into *č*. This is related to umlaut usually but not always (e.g. not in forms ending with *-in*), but even in terms of development there are diverse phenomena at work here. Nonetheless, in annotating currently all such examples are grouped together.

O/PAR/preglas-in-cč:

- Zato data poklicati **princova | prinčeva** vrstnika in prijatelja Rozenkranca in Gildenšterna, ki Hamleta dobro poznata, naj se z njim pogovorita in odkrijeta vzrok njegove norosti.

## 2.3 Additional Annotation

At the morphological level there are a few groups of tags available that are used in this system merely as additional annotation. This means that they are ascribed only when the basic contextual error tag already exists and an additional tag facilitates the merging of data according to different criteria or the ascription of additional (optional) information.

### 2.3.1 Morphological variants = O/DOD/variante

The question of morphological variability does not conceptually belong to any of the principle annotational levels discussed. At the formal level the correction *grada – gradu* is comparable with a paradigmatic correction of the sort *\*grado – gradu*; even if the goal is just to improve style by opting for the more neutral or more standard of 2 legitimate possibilities, the correction serves as useful feedback for the student. It is thus sensible to annotate corrections to variants, though it best to do so as an additional tag, so that they are decidedly separate and can be removed from the enumerated linguistic issues as desired.

Reference resources are used in determining or verifying the variability of forms – currently, in addition to the Sloleks lexicon, primarily *Slovenski Pravopis* and *SSKJ*. The variability of endings is shown in dictionaries by the qualifiers *in* and *tudi*, e.g. *grád<sup>2</sup>-ú* and *grád -a*, while Sloleks explicitly annotates variable forms and displays them in blue. If there is a discrepancy between sources, Sloleks is given precedence and the problem is highlighted. Those working on future annotation projects should take care to produce a list of all possible variable forms. A few examples that are currently annotated as such: *grada – gradu*; *želj – želja*, *straha – strahu*, *dneva – dne*, *bojo – bodo*; on the other hand this would not include the pair *razredje – razredi*, which is not identified as variable in the reference resources.

O/PAR/neglagolska-končnica + O/DOD/variante:

- Njegova zagnanost se kaže v tem, ko se je odločil pobegniti iz **grada | gradu** in premagati Valjhuna, med tem ko se njegova vdanost in ljubeznivost kažeta v njegovih privolitvi Bogomili.
- **Te | Ti** družini sta bili Capuletovi in Montegovi.

### 2.3.2 Mati-hči = O/DOD/besede-mati-hči

There are also some examples, determined a priori due to high frequency, that due to the atypical nature of syntactical samples cause orthographical difficulties. The examples in the current version of annotating are: *mati/hči* and *otrok*. Additional annotation of these examples is justified by the fact that the high frequency and diversity of the problems associated with these words indicate that teachers will likely prefer to approach them comprehensively (as whole paradigms as opposed to on a case-by-case basis). Future work could include developing methods for automatic identification of such problems. The tag *mati-hči* is thus used to annotate examples where regular case forms were substituted, just as for non-standard forms that are constructed for problematic roots. The substitutions *mater-mati* and *hčer-hči* and vice versa are extremely common, but there are other problems with case forms, e.g. *mati* instead of the proper genitive, dative, or instrumental forms, or *materjo* instead of the accusative. Among the non-standard forms there also occur e.g. *matero* and *hčero*.

O/KAT/sklon-drugo + O/DOD/besede-mati-hči:

- **Mater | Mati** je bila zelo zaščitniška do Hamleta.
- Ker če bi nehal raziskovat očetovo smrt, ga ljudje ne bi imeli za blaznega in Polonij bi pustil **hči | hčer** in Hamleta pri miru.

We do not ascribe to this group problems related to the (otherwise cognate) words *materin* or *hčerka*.

### 2.3.3 Otrok = O/DOD/besede-otrok

The tag *otrok* is used to annotate examples where regular case forms were substituted, just as for non-standard forms that are constructed for problematic roots. The substitution *otroci* – *otroki* is also common, along with *otrokih-otrocih*. Among the non-standard forms there also occur e.g. *otrocmi* and *otrokih*.

O/KAT/sklon-drugo + O/DOD/besede-otrok:

- To, dobro in slabo vidimo tudi v osnovnih šolah med **otroci | otroki**.

O/PAR/neglagolska-končnica + O/DOD/besede-otrok:

- Ko je pred cerkvijo doživel prvo ponižanje, sem videla, kako zelo podobni smo si z **otrocmi | otroki** iz tistega časa.

## 3 VOCABULARY

To this level we ascribe all corrections to vocabulary items that are primarily substitutions of a single word for another, either along with preservation of or change to the word order or the phrasal structure. For the theory on distinguishing morphological corrections from lexical and spelling ones cf. Chapter 8.

## 3.1 Nouns

### 3.1.1 Erroneous proper noun = B/SAM/napačno-lastno

This category contains examples where the teacher has corrected a proper noun, e.g. a personal or geographic one. In the majority of cases this concerns the substitution of an incorrect name, and more rarely the replacement of a name with paraphrasis or semantic expansion/reduction.

#### B/SAM/napačno-lastno:

- Takrat je **Lovrenc | Lovro** Kuhar, sedaj bolj znan po imenu Prežihov Voranc, napisal knjigo Samorastniki.
- **Kolonij | Klavdij** namreč izve, da Hamlet nekaj ve, o njegovem dejanju in ga poskuša odstraniti.
- To pa ni edini pobeg v tem tednu, na begu sta tudi dva dečka iz **Vrhnike | Jesenic**.

Whenever in addition to the substitution of a proper noun there occurs an expansion into a noun phrase, both pertinent tags are ascribed.

#### B/SAM/napačno-lastno + B/MEN/samostalnik-zveza:

- Tartuffe pa je dramska igra, ki je nastala v franciji v klasicizmu od **Baumarscha | avtorja Moliera**, a je bilo prepovedano, dokler ga ni kralj Ludvik XIV. vzel pod okrilje.

### 3.1.2 Proper and common name = B/SAM/lastno-občno

This category contains examples where a proper name is substituted by a common one that is semantically appropriate for the given context; here also go examples of a common noun substituted by a specific proper noun. These corrections often serve to avoid repetition, but this is not necessarily the case.

#### B/SAM/lastno-občno:

- **Hamlet | Princ** je bil res zaljubljen **v dekle | Ofelijo**, vendar ji ni zaupal.
- Govori o **neki družini | družini Khalan**, ki je bila dokaj bogata vendar se v njej niso mogli v miru pogovarjati.
- **Visoška kronika | Roman** opisuje dogajanje med Tridesetletno vojno, pa tja do sredine osemnajstega stoletja.

Whenever there is an example where the correction also targets phrasal expansion or reduction, the additional annotation is also ascribed.

#### B/SAM/lastno-občno + B/MEN/samostalnik-zveza:

- **Dramsko delo | Antigona** prikazuje boj za pravico, pri kateri tragično zmagata dobra stran
- Kralj organizira prijateljski spopad med **Hamletom in Laertom | mladeničema**.
- Kreon je optožil tudi **njeno sestro | Ismeno**, ki ni bila nič kriva.

### 3.1.3 Common vocabulary = B/SAM/občno-besedišče

This category encompasses substitutions of individual common nouns with another common noun. Often this includes corrections to semantically inappropriate or insufficiently specific nouns.

#### B/SAM/občno besedišče:

- To dejanje štejemo za vrh zato, ker se začne po njem vse razpletati, pa tudi zato, ker so izpostavljeni dve najbolj temeljni sporočili celotne **knjige | komedije**.
- Da bo počakal, da pridejo **pogani | kristjani** v svoje kočje in da jih morajo to noč doseči.

To corrections that arose because a certain vocabulary item is stigmatized (e.g. *deka – odeja*) we can further ascribe additional annotation for stigmatization (cf. the criteria in Chapter 3.10).

#### B/SAM/občno-besedišče + B/DOD/zaznamovano:

- Bil je obtolčen, pomankljiv **avtoček | avtomobilček**, ki ga nenadoma poboža, stisne v pest in ga odloži v torbico.
- Oseba, ki jo poznam je vztrajala, da ne bo naredila **faksa | fakultete** in ne bo imela željenega poklica.

## 3.2 Verbs

### 3.2.1 Verbal prefixes = B/GLAG/predpona

This category contains examples where the teacher has substituted a verb for a different verb with the same root but a different prefix.

#### B/GLAG/predpona:

- Tragedija govori o dekletu Antigoni, ki hoče **zakopati | pokopati** bratovo truplo, vendar ji kralj Kreon tega ne dovoli.
- Ona je še naprej bila mirna in je jokala, Metod je pa kasneje **odstopil | izstopil** z in si ogledal še zadnjič obraze potnikov skozi okno.
- Nato so se stvari začele **poboljševati | izboljševati**.

Here are also grouped those examples where there was a switch between a prefixed and unprefixed verb.

#### B/GLAG/predpona:

- Odloči se, da bo probleme **prevzel | vzel** v svoje roke in se maščeval za pravico.
- Spoznala sem, da moramo v življenju poskrbeti tudi, da se imamo sami lepo in ne misliti, kaj bodo drugi na to **porekli | rekli** in kako bo na njih vplivalo, kajti živi se samo enkrat.

This also includes variable prefixes or prefixes where the influence of colloquial usage is suspected (e.g. the group *iz-/z-*), which are treated as vocabulary issues and not spelling ones.

#### B/GLAG/predpona:

- Ko pa je **zvedel | izvedel**, da je Nora s pomočjo preprečila, da njeno dejanje vpliva na moža jo je sprejel nazaj, kot da se ni nič zgodilo.

If the corrected prefix changes the verbal aspect (perfective to imperfective or vice versa), the tag for verbal aspect is also attributed.

B/GLAG/predpona + O/KAT/vid:

- Ko sem se **tuširal | stuširal** sem šel v dnevno sobo.
- Drugače pa je bil družinski človek in vedno **pomislil | mislil** na druge potem pa na sebe.

In rare examples it is clear from the corrections that, in addition to the choice of verb, the teacher also found a morphological problem. These examples are also attributed an additional tag.

B/GLAG/predpona + O/PAR/glagolska-osnova:

- potem je risel font in **primil | poprijel** mojega brata za glavo in mu dojal glavo v voduv.

### 3.2.2 Substitution moči-morati = B/GLAG/moči-morati

These examples feature teachers changing the verb *moči to morati* or vice versa. This category also includes prefixed forms of *-moči* and *morati* (e.g. *premoči*).

B/GLAG/moči-morati:

- **Mogel | Moral** je pripluti mimo dveh siren, ki sta z svojima hipnotičnima glasovoma zvaljali mornarje na čeri.
- Občutka res ne bi **moral | mogel** opisati.
- Brž po prihodu sem ubil še Parisa, njenega zaročenca z vsem sovraštvom, ki sem ga **premoral | premogel**.

### 3.2.3 Other substitutions of modal verbs = B/GLAG/naklonski

This category includes corrections that substituted between various modal verbs, except for *moči-morati*, which, due to their frequency and specific nature, are grouped in a separate category. This one thus includes corrections made to: *hoteti*, *želeti*, *smeti*, etc. Here we also group corrections substituting a non-modal verb for a modal verb and vice versa.

B/GLAG/naklonski:

- **Hoče | Želi** tudi imeti zvezo s hčerko profesorja De Martinisa, a oba zaradi slabih izkušenj, ki sta jih v življenju doživljala, zveze ne moreta zaživeti.
- Čeprav je videl in slišal duha svojega očeta ne **želi | more** verjeti, da je to res.
- Prav tako je **moral | šel** v vojno tudi Berk, ki je svoje življenje v vojni začel s prepričanjem, da je vsak človek izrazito individualist, vojna pa ga je naučila, da je za preživetje nujno, da se človek bori za idejo skupnosti.
- Opravičil bi se ti, ker nisem **prišel | mogel priti** na tvoj kosanjev piknik, ki sem se ga ful veseliv.

Examples where the change occurs in changing a finite verbal form into the infinitive as well as in a modal verb (*prišel – mogel priti*) are annotated separately from examples where there

is a fundamental change at the phrasal level and are further additionally annotated with the appropriate annotation.

B/GLAG/naklonski + B/MEN/glagol-bz:

- Ves čas pa je **imel v mislih | želel** predstaviti svoj odnos do družbe in kakšne so bile, tako za kmeta, kot plemiča.
- Ko sem | se nekega dne **hotel pogovor | pogovarjal** z materjo, sem imel občutek, da se za zaveso skriva stric in prisluškuje, zgrabil sem meč in ga zabodel skozi zaveso.

Corrections concerning the addition or removal of the auxiliary particle *bi* are additionally annotated at the morphological level.

B/GLAG/naklonski + O/KAT/naklon:

- Želim si oditi v novi kraj, kjer **bi spoznala | si želim spoznati** ljudi, ki bi me razumeli in bi me pustili, da sem taka kot sem in ne da bi se morala pretvarjati, da sem nekaj drugega.

Whenever there is a correction to the verbal prefix, both appropriate tags are ascribed.

B/GLAG/naklonski + B/GLAG/predpona:

- Upam, da ne, in verjamem v to, da bom tudi jaz nekega dne srečala osebo, s katero bom **zmogla | mogla** preživeti življenje.

### 3.2.4 Other substitutions of verbs = B/GLAG/drugo

This category includes all single-word substitutions of verbs where the correction does not fit into any of the ones enumerated thus far. Often this includes corrections to semantically inappropriate or insufficiently specific verbs.

B/GLAG/drugo:

- Tudi razni krožki ki so jih vodili ugledni ljudje, so se počasi začeli zavedati slovenščine in so pri njej vse bolj **ustvarjali | vztrajali**.
- To lahko jasno vidimo s tem, kako je **napravit | ravnal** baron.

To corrections that clearly arose because a certain vocabulary item is stigmatized (e.g. *rabiti – potrebovati*) we can further ascribe additional annotation for stigmatization (cf. the criteria in Chapter 3.10).

B/GLAG/drugo + B/DOD/zaznamovano:

- Jim pokazati, da se je za svoj jezik potrebno boriti, saj ga za svoj obstoj **rabimo | potrebujemo**.
- Človek je najpametnejše bitje na svetu, to pomeni, da smo dovolj pametni, da se med seboj ne **kregamo | prepiramo**, da smo združeni.

For some corrected verbs it is difficult to assess whether the correction is the substitution of two words or a correction to one word that has been formed by false analogy (which would belong to the morphological level). If the original word can be found in the reference style guides, like SSKJ, it is annotated at the level of vocabulary, even if it is noted as archaic or rare in the dictionary.

B/GLAG/drugo:

- Zaradi velike sile je z levo roko **trknil | trčil** ob volan zaradi katerega si je zelo močno poškodoval mišico.

### 3.3 Pronoun

#### 3.3.1 Reflexive possessive = B/ZAIM/povratna-svojilnost

Here are gathered all corrections where the teacher corrects a possessive or personal pronoun (e.g. *njegov, njen, on, ona*) into a reflexive pronoun (*svoj, se*) or vice versa.

B/ZAIM/povratna-svojilnost:

- Morali bi znati odpuščati in ljubiti vse, tudi **naše | svoje** sovražnike.
- Nenazadnje mu še **svoja | njegova** lastna žena ne verjame.
- Ko sem zagreto razmišljala sem zraven **mene | sebe** zaslišala glas.

#### 3.3.2 Substitution ki – kateri =B/ZAIM/ki-kateri

This group features all examples where the teacher corrects *ki* into *kateri* or vice versa (including all case forms and prefixed extensions of *kateri*).

B/ZAIM/ki-kateri:

- Zelo velik udarec je tudi uničena kunčja Farma, **katera | ki** je bila eden redkih virov zaslužka za družino.
- Bolj bi tu povdarila Kajna, **ki | za katerega** se mi zdi, da ima zelo močno osebnost.

#### 3.3.3 Other problems with relative pronouns = B/ZAIM/oziralni

This group contains examples where the problem is in relativity: the teacher has annotated as incorrect the use of a relative pronoun or the relative pronoun has been substituted for another one. In categorizing we made sure to consider all forms of such pronouns, namely: *kar, kdor, kakršen, kolikšen, kateri* (except in combination with *ki*, which forms a separate group), *čigar*, and *kolikor*, etc.

B/ZAIM/oziralni:

- Napisal jo je že leta 1790, zaradi **česa | česar** velja danes za prvo pravo slovensko dramsko besedilo in s tem postalo pomembnejše od Škofjeloškega pasijona.
- Prav tako je bolši zrak, edina stvar **ki | česar** na kmetiji ni, je nočno življenje, zvečer po večerji se vsa družina odpravi spat.
- Živeti, tako dobro in slabo, trpeti in biti srečen, pasti in iti dalje, vse to je to kar nas oblikuje in **čemur | temu** pravimo življenje.



### 3.3.4 Substitution of negative pronouns = B/ZAIM/noben

This category groups various examples of substitutions with the negative pronouns: *noben*, *nobeden*, and *nihče*.

B/ZAIM/noben:

- Čeprav jim ponudi tudi predajo, ga **nobeden | nihče** ne zapusti.
- Zdravstveni delavci do pacientov nimajo tako nečloveškega odnosa, sploh pa **nobenemu | nikomur** ne odrežejo zdrave noge.
- Pripovedovalec je ugotovil tudi, da se **nobeden | noben** drug potnik ni zmenil za to gospo, in niti za to, kdo je iz sebe izpustil neprijeten vonj.

When along with a substitution of the pronoun *noben* there also occurs the deletion of a different word, only the pronomial substitution is annotated.

B/ZAIM/noben:

- Samo jaz te bom skoz čakala, ker me **noben fant | nihče** ne zanima.

### 3.3.5 Other pronomial substitutions = B/ZAIM/drugo

This category includes all other pronomial substitution based on rare (less typical for the corpus) corrections at the level of pronomial form or type. Problems corrected by the teacher can be syntactic or semantic in nature.

B/ZAIM/drugo:

- Duh zahteva maščevanje, vendar **ga | tega** nočem in nemorem storiti.
- Dandanes **vsi | nekateri** ki imajo več denarja ali imajo boljšo službo kot drugi z viška gledajo na njih kot da so nekaj manj vrednega.

To the morphological level we ascribe substitutions between stressed, clitic, and prepositional forms of the personal pronoun.

O/KAT/oblika-zaimka:

- Če **njega | ga** ne bi bilo, bi bil svet lepši, saj ne bi bilo vojn, ker se ljudje ne bi imeli za kaj bojevat.
- Ta prizor so otroci slabo razumeli saj **njim | jim** ni bilo jasno zakaj jočeta.

## 3.4 Preposition

### 3.4.1 Prepositions in verbal phrases = B/PRED/glag-zveze

This category groups examples where a non-modal verb is found in combination with a preposition that attaches to a non-verbal word or a phrase in an oblique case, e.g. *privoliti k*

*ideji – privoliti v idejo.*<sup>7</sup> The correction is annotated at the level of the preposition that the teacher corrects, making a more typical phrase in the process.

#### B/PRED/glagolske-zveze:

- Prešeren spozna kakšen je svet v resnici, stremi **k | za** idealom | ideali, kot je ljubezen do Primičeve Julije, a se zaveda, da so izven njegovega dosega.
- Kreona, Antigoninega očeta, pa težko uvrstimo **pod | med** strogo držljive ali prilagodljive ljudi.
- Prikazuje odnos med nebogljeno hčerko Lojzko, ki je bolehalo **s | za** hudo boleznijo na nogah in njenima staršema.

If a preposition is added or removed, the correction is attributed to the syntactic level.

#### S/IZPUST/predlog-drugo:

- Oče ga je sedaj vzel \_\_\_ | **s** sabo v Ljubljano, kjer ga je vzgajal s trdo roko.
- Postala sem jezna, da sem prišla na to velikanko zaradi trenerja \_\_\_ | **za** katerega sem mislila, da je smučar.

Whenever due to a substituted preposition leads to a morphological change elsewhere (e.g. *za pokop – pri pokopu*), and such a change is segmented as a separate correction, this change is annotated as a related correction.

To qualify for this group there must be present all three linguistic elements defined at the beginning. If there is no verb in the phrase, the correction does not go here. If there is no preposition, the correction is considered a problem of case at the morphological level. The same holds true if the student used the right preposition but the wrong case.

#### O/KAT/sklon-rt:

- V takem primeru bi ravnala tako, da bi človeka prepustila samega sebi in ko bi našel smisel življenja in priznal grehe, šele takrat bi si zaslužil **odpuščanja | odpuščanje** in **usmiljenja | usmiljenje**, ne le od mene, temveč tudi od Boga.

### 3.4.2 Prepositions into non-verbal phrases = B/PRED/neglag-zveze

This category groups all examples where the teacher corrected a preposition in a nominal, adjectival, or adverbial phrase composed of three constituents (*mnenje za probleme – mnenje o problemih*). The correction is annotated at the level of the preposition that the teacher corrects, making a more typical phrase in the process (if the teacher adds or removes a preposition, the correction is grouped at the level of syntax).

#### B/PRED/neglagolske-zveze:

---

<sup>7</sup> Without regard to the strength of the connectedness between the linguistic constituents in usage or the semantic transparency of the whole under consideration.

- Vsak človek ima svoje mnenje **za | o** takšne | takšnih probleme | problemih, kot so se dogajali v knjigi in se na žalost tudi v realnem svetu.
- Ob prekršku zakona **za | proti** pokop | pokopu storilca dejanja čaka smrt.
- Julija je bila jezna **do mene | name**, vendar ko je ugotovila, kakšno kazen mi je dal knez, je bila zelo žalostno.
- Bila je ena **izmed | od** 8 otrok, ki so se kasneje dobro razumeli.

Usually corrections to prepositions are followed by a change of case within the phrase. If such correction to the case is separately segmented, it is only grouped among related corrections.

### 3.4.3 Locative doublets = B/PRED/lokacijske-dvojnice

This category contains examples where the teacher substitutes prepositions that primarily denote location and come in pairs, and whose usage causes doubt: *iz-s/z* and *na-v*.

#### B/PRED/lokacijske-dvojnice:

- **V | Na** podeželju pa se poročajo veliko in se med seboj lepo razumejo da družina lahko šteje pet do šest otrok in ljudje so zadovoljni z delom in z drugimi stvarmi.
- Šli so na izlet **na | v** Rogaško Slatino.
- Tu se Črtomir izkaže za taktičnega, saj namerava ponoči zbežati **iz | z** gradu.

Here we group only those examples that describe a specific physical location. Less common are examples with metaphorical meaning that are typically annotated as corrections of a preposition in a verbal phrase.

#### B/PRED/glagolske-zveze:

- A naletijo **v | na** začaran krog, ker se bo nekdo nazaj njim maščeval in temu ni konca.
- Če pa pogledamo **iz | z** druge strani, je Bog s tem Kajna kaznoval, saj mora nositi to breme, da je ubil svojega lastnega brata.

### 3.4.4 Other substitutions of prepositions = B/PRED/drugo

Here are grouped examples that did not fit into any of the above defined category. The rule here applies that potential corrections to the case of the nominal phrase following the correction are annotated as related corrections at the level of morphology. If some other category is also corrected or a word replaced, that is also appropriately annotated.

#### B/PRED/drugo:

- **Na | Med** svojem vojaškem potovanju, v borbi za preživetje, je bolje spoznal soborca Antona.
- Visoki stanovanjski bloki, športne hale, mejni prehodi in drugi objekti kakor v Sloveniji tako **po | v** Rusiji, Nemčiji, Italiji in v Avstriji kjer je zdaj zaposlen.
- Podobna tema je tudi **pri | v** Shakespeareovi drami Romeo in Julija, kjer zaradi sovraštva oba umreta.
- Če želimo biti pravični in usmiljeni, se moramo ravnati po svoji vesti, ne smemo pa dopustiti, da nam usmiljenje zamrači um **pred | za ceno** pravičnostjo | pravičnosti.

To this category we ascribe substitutions of prepositional nominal phrases if they serve a similar function and similar meaning.

B/PRED/drugo:

- V obeh delih se zrcali želja po oblasti, pri Antigoni si jo želi Kreon pridobiti **skozi | s pomočjo** zaupanje ljudstva, komunistični režim pa si jo utrjuje s prisilnim spreminjanjem ljudi, ki se bi mu lahko uprli.
- Omenjeno besedilo, ki predstavlja kot težavo izdajstvo bližnjih, govori o zgodbi Hamleta, katerega oče je bil umorjen **zaradi | s strani** njegovega lastnega brata Klavdija.

### 3.5 Conjunction

#### 3.5.1 Substitutions of *in-pa-ter* = B/VEZ/*in-pa-ter*

This category contains examples where the teacher variously corrects the coordinating conjunctions *in*, *pa*, and *ter*.

B/VEZ/*in-pa-ter*:

- V njegovi zgodbi se Jerobnik preseli s podeželja v Trst in se preimenuje **in | ter** popolnoma asimilira v italijansko okolje iz koristi.
- Bog je to videl in Kajna izgnal iz raja ter zapovedal, da bo do smrti taval naokoli **ter | in**, če ga bo kdo ubil bo sedem krat maščevan.
- Imamo primere, kot so Janez **pa | in** Meta.

#### 3.5.2 Coordinating adversative conjunction = B/VEZ/*protivni*

This category contains examples where the teacher corrects one coordinating adversative conjunction with another: *a*, *toda*, *vendar*, *pa*, *ampak*, *temveč*.

B/VEZ/*protivni*:

- Tudi v današnjem času, bi bili potrebni še enega razsvetljenstva, saj ne uporabljamo več svoje pameti, **vendar | ampak** se ravnamo po modi, po tistem, kar drugi pravijo.
- Delal se je zelo vesel, da sta prišla, **ampak | vendar** jima ni zaupal.
- Ljubezen v renesansi ni samo ljubezen med fantom in dekletom, **vendar | temveč** tudi ljubezen do staršev, kot je v primeru zgodbe o Hamletu.

This category also includes the particle *samo* when used as a conjunction.

B/VEZ/*protivni*:

- Tu se Bog pokaže v drugi luči in se Kajna usmili, ga pusti pri življenju, ga ne kaznuje s težkim delom, ponudi mu celo zaščito, **samo | a** vedno po taval zgubljen po svetu.

#### 3.5.3 Change to subordination = B/VEZ/*sprememba-odnosa*

This category features examples where the teacher's substitution of a conjunction changes a clausal relationship from subordinate to coordinate or vice versa.

B/VEZ/*sprememba-odnosa*:

- Ravnal je po človeški pravici, saj je bil čisto po pravici užaljen, **saj | da** je Bog brez razloga zavrnil njegovo daritev.
- Maščevati bi se morali tako, da bi se krivec iz maščevanja nekaj naučil, **da | ter** ne bi več počel takih grozot.

### 3.5.4 Other substitutions of conjunctions B/VEZ/drugo

This catch-all category contains corrections to conjunctions that do not fit in any of the above-mentioned categories. Often here are found corrections to conjunctive clauses or clauses that function like conjunctions, changing them into one-word prepositions or vice versa.

B/VEZ/drugo:

- Vsebuje tudi dialog, **kjer | ker** ima dramske sestavine.
- Za 2. Domače branje sem si izbral knjigo z naslovom Punce za znoret, saj zelo rad berem dela, ki so vezane na mladino **in sicer | torej** na nas.
- Bil je le eden izmed mnogih, ki je bil po mnenju države nevaren **a kljub temu | čeprav** ni imel nekega visokega položaja v družbi.
- Temeljna resnica, ki pa Berka najbolj boli je, da v vojni preživiš le če se naučiš ubijati, **ker če ne | sicer** si sam ubit.

### 3.6 Adjective = B/PRID/drugo

These examples show where teachers have corrected an adjective into a different one. These corrections are usually at the level of changed meaning or more precise semantic definition.

B/PRID/drugo:

- Za zaključek bi lahko neskončne ure pisali, kaj je prav, **trda | trdna** načela ali prilagajanje.
- K temu jo žene **bratovska | sestrska** ljubezen in pa seveda splošno človeško nagnjenje, da je treba mrtve pokopati.
- Bil sem besen | jezen!

Problems where a definite possessive adjective formed from a personal name according to the paradigm for the wrong gender are grouped at the level of morphology.

O/KAT/spol:

- **Antigonovo | Antigonino** dejanje se Kreonu ni zdelo prav nič lepo.

Too corrections that arose because a certain vocabulary item is stylistically marked (e.g. *brihten – pameten*) we can further ascribe additional annotation (cf. the criteria in Chapter 3.10).

B/PRIDEVNIK + B/DOD/zaznamovano:

- Pri Matičku vidimo kako **brihten | pameten** je bil in se je zavedal kaj vse bi lahko dosegel če ne bi bil kmet.

Whenever there is a substitution concerning a multi-word adjectival/adverbial phrase, the corrections are grouped in the category for substitutions of various phrases in combination with an adjective or adverb.

B/MEN/pridevnik-prislov-bz:

- To se ni zdelo kaj preveč **modro | modra odločitev**, vendar pa nam to kaže Salomonovo moč, avtoriteto, pa tudi to, da si je upal svoj sloves postaviti na kocko, kajti nenavadno je sekati otroka na pol.

### 3.7 Adverb = B/PRISL/drugo

These examples show where teachers have corrected an adverb into a different one. These corrections are usually at the level of changed meaning or more precise semantic definition.

B/PRISL/drugo:

- Ali bo prav ali ne, boš videl **vnaprej | kasneje**.
- "Povej kje si bil!" je **nazaj | zopet** zakričala jezna Micka, a Marko je mirno vzel mleko in si ga začel nalivati v skodelico.
- Skoraj **zmerom | zmeraj**, če se prepirava ga oče zapre, ker mu gre na živce, da se skozi kregava za računalnik.

To corrections that arose because a certain vocabulary item is stylistically marked (e.g. *gluh – ravno*) we can further ascribe additional annotation (cf. the criteria in Chapter 3.10).

B/PRISL/drugo + B/DOD/zaznamovano:

- Z njim je imela tudi spolne odnose, kar za sedemnajstletnico ni **gluh | ravno** najbolj primerno.

### 3.8 Other parts of speech = B/OST/drugo

This heterogeneous group features collected examples where the teacher has made a correction because the student chose the wrong part of speech. There is an insufficient number of such examples so as to facilitate the establishment of independent categories, e.g. particle or quantifier. It is typical of these examples that the substitutions always remain in the same part of speech.

B/OST/drugo:

- Zdi se mi pravilno kot je ravnal on, **edino | le** prepozno se je odločil maščevati očetovo smrt, saj je s tem povzročil smrt veliko njemu ljubih oseb.
- Pred gosti sta starša igrala idealen par in se ves čas nasmihala **eden | drug** drugemu, vendar v resnici ni bilo tako in prav zaradi tega sta se Lojzki gnusila.

### 3.9 Substitutions beyond the confines of part of speech

This category contains collected examples where a word within a specific part of speech is changed to a word from a different part of speech or a single word is replaced by a word phrase.

#### 3.9.1 Lexical word or phrase changed into pronoun = B/MEN/polnopomenska-v-zaimek

This category contains collected examples where the teacher replaces a lexical word with a (personal, demonstrative, or possessive) pronoun, as well as substitutions with the word *leta*.

B/MEN/polnopomenska-v-zaimek:

- Niso mogli postati njegovi pravi prijatelji, ker jim **Hamlet | on** ni zaupal, kar pa mu je bilo pri prijateljih najbolj pomembno.
- Našla **je kuhinjo | jo je** in začela kuhati.
- **Kupec | Ta** mu ni verjel in ga je šel prijaviti na policijo.

In this category are grouped those corrections where a pronoun substitutes a phrase.

B/MEN/polnopomenska-v-zaimek:

- Prvi razlog **za njegovo blaznost | zanjo** je bila po kraljevem mnenju smrt Hamletovega očeta, drugi pa njegova ljubezen do Ofelije.

#### 3.9.2 Pronoun to a lexical word or phrase = B/MEN/zaimek-v-polnopomensko

This category contains collected examples where the teacher substitutes a (personal, demonstrative, or possessive) pronoun with a lexical word or phrase that has a nominal or adjectival core.

B/MEN/zaimek-v-polnopomensko:

- Zato je Julijin oče našel **njej | hčeri** snubca, za katerega bi se naj poročila, zato je to večer pripravil zabavo v njihovem imenu.
- Za Majhne ljudi, kamor spada tudi **on | sam**, pravi, da so le statisti v tej veliki igri.
- V zgodbi, kot je **ta | Tržačan** je najbolj žalosten dogodek smrt, a smrt ne pride na plano le v te zgodbi, poznam zgodbo katere konec je dečkova smrt.

#### 3.9.3 Substitution of conjunction and pronoun = B/MEN/veznik-zaimek

To this category are ascribed examples where the teacher corrects a pronoun into a conjunction or vice versa. This category primarily contains examples with substitutions with the words *ki* and *in*. The literature is inconsistent in terms of how to classify the words *ki* and *kar*, but for the purposes of consistency we interpret them in all examples as pronouns<sup>8</sup>

---

<sup>8</sup>In harmonizing our categories of parts of speech it made sense to make some reclassifications, as these corrections target the composition of sentence-level relationships that are of the same nature as other already extant categories.

#### B/MEN/veznik-zaimek:

- Takrat se opogumi tudi baronica **ki | in** prispeva svoj del s tem ko zamenja obleke z Nežkinimi.
- Sam tisti **ko | ki** so revni se jim tudi neda delati raj vlečejo denar od borze kot da bi sli delati.

#### 3.9.4 Word family = B/MEN/besedna-družina

This category contains corrections that transcend parts of speech but remain within the same word family. Typical examples include substitutions of nouns and adjectives (*Abel – Abelov; pogum – pogumen*), nouns and verbs (*skrb – skrbeti*), and adjectives and adverbs (*drugačno – drugače*).

#### B/MEN/besedna-družina:

- saj je bil **pogum | pogumen** in zelo razumen ter močan.
- Opravljati so morale dela kot so gospodinjstvo, **skrb | skrbeti** za otroke in ogenj.
- Danes je to povsem **drugačno | drugače**.

There are also rare examples where the teacher replaces a word either written in a non-standard orthography or with other spelling issues, but it is clear that the substitution changes part of speech (therefore it is not merely a spelling correction). This is annotated as a corrected vocabulary item, but it is grouped in the “other” category.

#### B/MEN/drugo:

- Še posebej zaključek mi všeč, ko nakoncu Herod naroči tudi **ubo | ubiti** Salome.
- Rekel sem če mi daš 1 dan časa za **razmiselt | razmislek**.

#### 3.9.5 Noun and phrase = B/MEN/samostalnik-bz

This category contains collected examples where the teacher corrects a noun, expanding it into a broader nominal phrase, or vice versa. Among the most common types of phrases at work here are ones that contain a matching adjective (*čas – časovno obdobje*) or pronoun (*drama – to besedilo*), noun in the genitive (*maščevanje – iskanje pravice*), prepositional phrase (*individualnost – zavest o individualnosti; duh – v podobi duha*), and similar. In the majority of examples the teacher has preserved the meaning, which the correction either expands or specifies semantically. Nonetheless, we also group here those examples where the correction more significantly alters the meaning.

#### B/MEN/samostalnik-zveza:

- Le kdo bi odrekel pomoč očetu, četudi je samo **duh | v podobi duha?**
- Oni pripravijo **kupe denarja | denar** in že imajo prostost.
- Odloči se, da bo probleme prevzel v svoje roke in se maščeval za **pravico | očetovo smrt**.
- V knjigi se pojavljata dva **časa | časovni obdobji** in kraja dogajanja.

Whenever the correction also features the substitution of proper nouns and regular ones, it is appropriately annotated as such.



#### B/SAM/lastno-občno + B/MEN/samostalnik-zveza:

- **Dramsko delo | Antigona** prikazuje boj za pravico, pri kateri tragično zmagaja dobra stran
- Kralj organizira prijateljski spopad med **Hamletom in Laertom | mladeničema**.
- Kreon je optožil tudi **njeno sestro | Ismeno**, ki ni bila nič kriva.

#### 3.9.6 Verb and phrase = B/MEN/glagol-bz

This category contains collected examples where the teacher corrects a verb, expanding it into a broader verbal phrase, or vice versa. Among the most common phrases that occur in these examples are ones with nouns (e.g. *imel mnenje – menil; imel v mislih – želel*), adverbs (*psihično uničuje – izčrpuje; stopi ven – izstopi*), infinitives (*zakopal – pustil ležati*), etc. In the majority of examples the teacher has preserved the meaning, which the correction either expands or specifies semantically. Nonetheless, we also group here those examples where the correction more significantly alters the meaning.

#### B/MEN/glagol-bz:

- Zato se takrat **razosebi | odpove čustvom/človečnosti** v sebi, čeprav mu je pomemben ...
- Pobral je ves denar, truplo **zakopal | pustil ležati** in odšel.

Attention must be paid in examples with modal verbs. Examples where the correction targets changing a finite verbal form into the infinitive as well as in a modal verb (*prišel – mogel priti*) are put into this category if there is a fundamental change at the phrasal level and are further additionally annotated with the appropriate annotation.

#### B/GLAG/naklonski + B/MEN/glagol-bz:

- Ves čas pa je **imel v mislih | želel** predstaviti svoj odnos do družbe in kakšne so bile, tako za kmeta, kot plemiča.
- Ko sem | se nekega dne **hotel pogovor | pogovarjal** z materjo, sem imel občutek, da se za zaveso skriva stric in prisluškuje, zgrabil sem meč in ga zabodel skozi zaveso.

#### 3.9.7 Adverb/adjective and phrase = B/MEN/prislov-pridevnik-bz

This category contains collected examples where the teacher corrects an adverb or adjective, expanding it into a broader phrase, or vice versa. Among the most common phrases that occur in these examples are ones with nouns (*dolgotrajno – dolgo časa; vdano – iz vdanosti; enakopravno – po tem zgledu; drugače – drugačna pot; zdaj – ta trenutek*). Often, thus, we are not concerned with adverbial phrases (i.e. they do not contain actual adverbs), but with phrases that typically function as adverbial qualifiers, sometimes also as other sentence-level constituents (e.g. *zdi se modro – zdi se modra odločitev*). To this category are also ascribed examples where the correction or corrected part contains a deictic demonstrative adverb (*tukaj – v tem primeru; tam – v njej, tako – v povedi; eno leto nazaj – pred enim letom*):

Corrections from an adjective to a phrase or vice versa are less common. There occur patterns here with adverbs (*hermetičen – hermetično zaprt*), nouns (*spoštljiv – vreden spoštovanja*), and similar.

#### B/MEN/prislov-pridevnik-bz:

- Moje mnenje o televizije je to, da lahko jo gledaš kot sprostitiv v prostem času ampak ne **dolgotrajno | dolgo časa** gledaš saj res škodijo očesu.
- Njihovo delovanje pa v tej komediji kaže popolnoma **drugače | drugačno podobo**.
- **Že v mladih letih | Pravkar** mu je umrla žena, zato se preseli v Trst, kjer uči v šoli.
- Namreč ko smo bili **eno leto nazaj | pred enim letom** v Barceloni sva se z mamo morali prilagajati samo njej.
- Zakaj bi se voznik vstavil, ter odprl vrata **hermetičnega | hermetično zaprtega** vozila, če je lako potem tudi on ogrožen.

#### 3.9.8 Other types of substitutions =B/MEN/drugo

To this category are ascribed all other corrections where there is a substitution of several part of speech and phrases. Similarly here are grouped combinations with functional units (preposition, adverb, conjunction, etc.)

#### B/MEN/drugo:

- V primeru komandanta Gregorja bi nasprotoval oz. **saj | vsaj** pomagal Katarini pri tem da jo ne ubijejo oz. **saj | vsaj** jaz je nebi ubil osebno.
- Njun najpomembnejši podvig je, ko skujeta zaroto proti baronu, pri čemer se baronica preobleče v Nežko, Nežka pa **obratno | vanjo**.
- Moje mnenje je, da je bil Hamlet zelo dober in zvest prijatelj, če si bil tudi ti **NJEMU | tak do njega**.
- **Tone partljič | Danes** je star 69. let.

### 3.10 Additional Annotation

#### 3.10.1 Stylistically marked vocabulary = B/DOD/zaznamovano

The tag for stylistically marked vocabulary items is used as additional annotation to corrections of lexical words. We add it when the teacher corrects a) a dictionary word qualified in the reference literature the qualifier dialectal, expressive, archaic, forbidden, colloquial, etc., or b) when the word is not found in the reference literature but it is still quite clear that it is a dialectal or jargon word.

#### B/SAM/občno-besedišče + B/DOD/zaznamovano:

- Oseba, ki jo poznam je vztrajala, da ne bo naredila **faksa | fakultete** in ne bo imela željenega poklica.

#### B/GLAG/drugo + B/DOD/zaznamovano:

- Jim pokazati, da se je za svoj jezik potrebno boriti, saj ga za svoj obstoj **rabimo | potrebujemo**.

#### B/PRID/drugo + B/DOD/zaznamovano:

- Pri Matičku vidimo kako **brihten | pameten** je bil in se je zavedal kaj vse bi lahko dosegel če ne bi bil kmet.

B/PRISL/drugo + B/DOD/zaznamovano:

Fringe cases include examples where the corrected word has been “Slovenized”, but it is not specially annotated with a qualifier in the dictionary. These examples are not ascribed the addition of stylistically marked vocabulary. The word *plan*, e.g., is marked in SSKJ as *journalistic*, but not in the context in which the student used it. In general we adhere to the principle that we tag only those examples where the stylistic markedness is quite obvious.

B/SAM/občno-besedišče:

- Gre celo do te mere, da da ubiti stražarja, ki poznata Antigoin **plan | načrt**.

## 4 SYNTAX

At this level we tag those corrections that target phrase-, predicate-, and sentence-level syntax, e.g. corrections to syntactic structures, word order, inter-sentence relationships, etc. A caveat is in order, namely that the current methodology of segmenting and annotating corrections in the Šolar corpus is not ideal for the syntactical level. It is thus expected that there will be relatively more conundra and lapsus in annotation at this level, and the decisions must be viewed through a different tool, and at the same time errors and corrections must be segmented if possible.

### 4.1 Word order

This category includes corrections where the teacher moves around a word or several words within the sentence. Examples where the teacher also changes, adds, or deletes a word are annotated with additional or other categories. Fundamentally speaking, corrections are divided into groups where a clausal constituent, individual word, or entire clause is moved. Whenever this division is not determinable with one category, the correction is grouped by analogy (which category is the most sensible?). In annotating we are interested only in the shift, not the reason for it, too (e.g. examples that are the result of functional sentence perspectives are not distinguished from grammatical corrections).

#### 4.1.1 Constituent order: predicate-subject = S/BR/povedek-osebek

In these examples the teacher corrected the word order of the subject and the predicate. The subject and predicate can be comprised of several words. This category also gets examples where the shifts are related to the partite nature of the predicate, e.g. when the subject encompasses two parts of the predicate that, given the correction, should remain together.

S/BR/povedek-osebek:

- Končno se je odločil, da je Antigona delovala pravilno in se je **on zmotil | zmotil on**.
- Umr! je Polonij, ki ga je Hamlet zabodel | zabodel Hamlet.
- Dedek in mama sta se seveda precej napila, zato je **vožnja domov z avtom imela | imela vožnja domov z avtom** posledice.

#### 4.1.2 Constituent order: povedek-predmet = S/BR/povedek-predmet

In these examples the teacher corrected the word order of the object and the predicate. The predicate and subject can consist of several words.

S/BR/povedek-predmet:

- Don Kihot pa za razliko od Hamleta **veliko stvari dela | dela veliko stvari**, ki so pa večinoma nepravilne in neumne poteze.
- Skleneta, da bosta s prevaro **Hamleta ubila | ubila Hamleta**, ubojstvo pa naj bi ispadlo le kot nesrečno naključje.

To this category we also ascribe examples where the correction contains altered order of the predicate and two objects in different cases.

S/BR/povedek-predmet:

- Njegov pomočnik pa ker misli, da je use res, **odreže nogo Vebri | Vebri odreže nogo**.

The predicative modifier is considered part of the predicate, wherefore such examples are also ascribed here. Similarly we also group here examples where the teacher switches the order of the object and the predicator's determiners.

S/BR/povedek-predmet:

- In ta kazen je, da jo zaprejo v kamnit grob živo | živo zaprejo v kamnit grob.
- Berk se spominja vojne kot lepe izkušnje, ampak prepričan je da **generalove ukaze je potrebno | je generalove ukaze potrebno** spoštovati in jih izpolnjevati.

Examples where clitics serve the role of an object, which the teacher moves due to its obligatory syntactic second position, are grouped among corrections to clitic strings.

S/BR/naslonski-niz-znotraj:

- Bogomila pa pravi da bo ostala zvesta Bogu in Črtomirju in **čakala ga bo | ga bo čakala** v nebesih.

#### 4.1.3 Order: sentence-adverbial determiner = S/BR/povedek-prislovno-določilo

In these examples the teacher corrected the word order of the object and adverbial determiner. To this subcategory are ascribed examples where the teacher switches the order of exactly two clausal constituents – the predicate and the adverbial determiners, which can be composed of several words. This category also gets examples where the shifts are related to the very partite nature of the predicate, e.g. when the adverbial determiner encompasses two parts of the predicate that should remain together.

S/BR/povedek-prislovno-določilo:

- Ljubizen je v vsakemu od nas, nekateri jo močneje začutimo in **pokažemo | začutimo** in pokažemo močneje kot drugi.
- Ko je Tonček **domov prijokal | prijokal domov**, ga je mati vprašala kaj je narobe.

#### 4.1.4 Order: particle = S/BR/členek

To this category we ascribe examples where the word order is switched because the teacher moves a particle or a particle phrase within a clause, including discursive elements such as *denimo*, *na primer*, *recimo*, etc. These examples are grouped under substituted clausal constituents through analogy with substitution of adverbial determiners, even though linguistically speaking a case could be made for other categorization.

##### S/BR/členek:

- V želji, da bi Jermama zlomil, **celo omeni | omeni celo** njegovo mater.
- Če se že ustavimo **pri šoli na primer | na primer pri šoli**, se velikokrat zgodi, da so učitelji oz. profesorji do nas usmiljeni, čeprav bi se morali ravnati po šolskih pravilih.

#### 4.1.5 Order within clausal constituents = S/BR/znotraj-stavčnega-člena

In these examples the teacher changed the word order within an individual clausal constituent. Here we also group examples where the word order was substituted within the predicate (e.g. the position of the verbal attribute and the determiner).

##### S/BR/znotraj-stavčnega-člena:

- Hamlet se, ko začne pripovedovati o **resnici in lažeh | lažeh in resnici**, ki so jih skrivali pred njim, izredno osami.
- Od velike žalosti se ubije tudi **sin Kreona | Kreona sin** in Antigonin zaročenec.
- Pa še **ne zna razmišljat | razmišljati ne zna** dovolj, saj ga drugi vedno ukanijo, kljub temu, da hoče on njih.

#### 4.1.6 Clitic string: order of clitics = S/BR/naslonski-niz-znotraj

Problems concerning the word order of clitics are classified as such on a priority basis. The subcategory of “order of clitics” typically receives examples where incorrect word order arises within a clitic string due to formation of the conditional or the future, as well as in combination with *naj*.

##### S/BR/naslonski-niz-znotraj:

- Na dan, ko **se bi | bi se** morala Julija poročiti, je odhitela do Lorenza po napoj, da bi bila 48 ur mrtva.
- Po takšnem mišljenju **se bi jaz | bi se jaz** odločil za maščevanje.
- Torej podala sem se po hribu navzdol, nevedoč kaj vse **se bi mi | bi se mi** lahko pripetilo.

#### 4.1.7 Clitic string: independent-subordinate = S/BR/naslonski-niz-prirednost-podrednost

These are examples where the teacher has changed the position of the clitic string or individual clitics because their structure in the independent or subordinate clauses differs. These examples are distinguished by the fact that the clitic string usually occurs near the subordinating or correlating conjunction, as well as by the fact that the structure would still be considered well-formed in a different intraclausal arrangement.

S/BR/naslonski-niz-prirednost-podrednost:

- Kar se pa identitete tiče lahko povem samo to, da **nekoč nam jo je zatiral | nam jo je nekoč zatiral** sistem, danes pa jo izgubljam sami zaradi svobode, ki jo imamo in liberalnosti družbe.
- V tem, kakor tudi v vseh drugih razmerjih, bi se po mojem mnenju morali sklepati kompromisi in **bi si partnerji | partnerji bi si** morali pustiti dovolj prostora in možnosti za uresničevanje lastnih sanj, saj žrtvovanje za nekoga drugega prinese samo notranjo bolečino in zagrenjenost.

#### 4.1.8 Other changes to word order = S/BR/drugo

In this group are collected examples where the teacher corrects word order, but there is no substitution as in the listed subcategories. Here are grouped more complex and less typical examples.

S/BR/drugo:

- Stara je enako dvanajst let, **ima lepo rjavkasto obarvano barvo**, oblečena je v bela oblačila, dražja, iz Italije, ter dolge blond lasje lase **in lepe modre oči**. | **Ima lepo rjavkasto obarvano barvo in lepe modre oči**, oblečena je v bela oblačila, dražja, iz Italije, ter dolge blond lase.
- Za primer bom vzel naš razred, ki se po mojem mnenju ne drži skupaj in vsak preveč gleda samo nase, še posebej, če **koga prosi kdo | kdo koga prosi** se ta le izgovori, da ne zna snovi ali pa reče, da nima časa.

It is characteristic of these examples that they often transcend clausal borders. Here are grouped, e.g., examples where the teacher moves a clausal constituent from the main clause to a subordinate one or vice versa.

S/BR/drugo:

- **V boju je odločen, da bo zmagal | Odločen je, da bo zmagal v boju** in dosegel svoj ideal — ohranil bo vero svojih staršev in pridobil domovino, v kateri bodo ljudje svobodni.
- **Na Gregorjevem mestu dvomim, da bi | Dvomim, da bi na Gregorjevem mestu** zbrala toliko poguma, ter pritisnila na petelin pištolo, ki bi ubila mojega ljubljenega, moj smisel življenja, ki ne bi bil krvav kot vojna.

Also grouped here are examples where the structure of embedded clauses or phrases is changed, along with less frequent examples where the teacher changes the order of entire clauses.

S/BR/drugo:

- V satirični drami **pisatelja Moderne, Ivana Cankarja, Hlapci | Hlapci pisatelja Moderne, Ivana Cankarja**, je konflikt med legitimnim in legalnim predstavljen, kot konflikt med Jermanom in klerikalno oblastjo.
- Že naslednji dan sledi sojenje, v katerem obtoženega spoznajo za krivega, saj menijo, da če ga ne gane niti materina smrt je zagotovo pokvarjen in sovražen družbi | je zagotovo pokvarjen in sovražen družbi če ga ne gane niti materina smrt.

- Verjamem pa v to, da vse kar si kdajkoli storil slabega se ti bo | se ti bo vse, kar si kdajkoli storil slabega, vrnilo v enaki količini, samo vprašanje časa je kdaj!

To this category are ascribed the otherwise rare examples where the teacher corrects an emphasized subject, as demonstrated by the first example below, and, lastly, examples where the word is problematic in a textual sense, e.g. when filling out forms, writing addresses, etc.

S/BR/drugo:

- **Ožbej, ko se je | Ko se je Ožbej** preveč napil in bil omotičen je začel čutiti neko mehko v sebi.
- Še enkrat se vam zahvaljujem. XImeX XPriimekX XImeX XPriimekX **7 . 6. 2010, XKrajX | XKrajX, 7. 6. 2010**

## 4.2 Omitted constituents

This category contains examples where the teacher adds an omitted word or phrase. These categories are divided by part of speech and other characteristics of the primary added word. In the current categorization we do not distinguish between adding grammatically necessary elements and those that merely flesh out or edify a text's content.

### 4.2.1 Noun: common noun = S/IZPUST/samostalnik-občno-ime

In these examples the teacher has added a common noun to the clause, potentially with an appropriate preposition, or a simpler phrase with a common noun at its core (regardless of whether the added element in the whole clause forms part of a proper noun or not, e.g. in the example below).

S/IZPUST/samostalnik-občno-ime:

- Ker je Julijo skrbelo, da ne bi \_\_ | **starši** izvedeli da sta se poročila je Julija vprašala meniha, če ji lahko pomaga.
- Nepravično pa se mi je zdelo \_\_ | **ravnanje očeta** za starejšega sina.
- Antigona je umrla zaradi tega, ker je kralj dajal prednost državnim zakonom, ne pa božjim, Simon pa je zaradi bolanega strahu države pred revolucijo, izgubil nogo \_\_ | **in zdrav razum.**
- Zato so jo leta XLETOX sprejeli v Črnovrški klub \_\_ | **gorskih kolesarjev.**

In ascribing examples here we follow the word order at the heart of the phrase, wherefore we also encompass various examples of phrases that function as adverbial determiners:

S/IZPUST/samostalnik-občno-ime:

- Polikarp je bil \_\_ | **svoje dni** vojak.

### 4.2.2 Noun: proper noun = S/IZPUST/samostalnik-lastno-ime

In these examples the teacher has added a proper noun to the clause, potentially with an appropriate preposition, or a simpler phrase with a proper noun at its core.

#### S/IZPUST/samostalnik-lastno-ime:

- To je bilo takrat, ko sta \_\_ | **z Matevžem** odšla k gospodu Andreju po tržaske dinarje
- V tej knjigi \_\_ | **Tavčar** opisuje, da je bil Janez odvetnik in je živel v Ljubljani.
- V romanu Senčni ples pa ne spremljamo vojnega dogajanja, ampak obdobje med leti 1951 in 1954, ko še ni bila določena meja \_\_ | **med Jugoslavijo in Italijo.**

Often there are borderline examples, where the teacher has added to the sentence the name or author of a written work, often together with a lay description of the genre. Such examples are grouped under proper nouns.

#### S/IZPUST/samostalnik-lastno-ime:

- Berk \_\_ | **v romanu Menuet za kitaro** je partizan, želi si boja, dokler ga še ni okušal .
- Antigonin pokop brata je osrednji dogodek \_\_ | **Sofoklejeve tragedije Antigona.**

#### 4.2.3 The verb biti = S/IZPUST/glagol-biti

In these examples the teacher added to the sentence an omitted instance of the verb *biti* in various forms and roles, with the most typical being an omission in the formation of the past tense.

#### S/IZPUST/glagol-biti:

- Jezus je seveda pozitivna oseba, to je razvidno iz celotne zgodbe, poleg tega pa je on tudi edini ki te grešnice ni obsojal, ampak ji \_\_ | **je** dal novo priložnost da ravna pravilno.
- Včasih se izkaže, tako jaz menim, da \_\_ | **je** če se preveč strogo držiš svojih načel lahko še slabše.
- Družina se je \_\_ | **je** sramovala.

Whenever the added verb triggers a change to tense, the correction is annotated at the level of morphology.

#### O/KAT/čas:

- Na plaži pride do pretepa, in Raymond je \_\_ | **bil** ranjen.

#### 4.2.4 Other omitted verbs = S/IZPUST/glagol-drugo

In these examples the teacher added to the sentence an omitted lexical verb. In more rare examples also a phrase has been added to flesh out the meaning. Since in segmentation such examples are treated as just one correction, they are categorized as such whenever doing so makes sense given the original omission. If, e.g., it is clear from context that the student perhaps omitted the part added to the clause in order avoid repetition, such a correction is grouped here.

#### S/IZPUST/glagol-drugo:

- Proti njemu nimam kaj \_\_ | **reči** in ga imam v družini najraje.
- Umori so se \_\_ | **dogajali**, se še \_\_ | **dogajajo**, in se še bodo dogajali — to je dejstvo.
- Imeli so ga za rešitelja, ko so vsi drugi \_\_ | obupali nad njim včasih pa že celo sami obupali nad sabo.



More complex additions, ones that can be interpreted as adding an entire clause, are handled elsewhere.

S/IZPUST/stavek:

- Starša bi se morala bolj posvetiti otroku \_\_ | **in se vprašati**, kaj je za njo bolje, da se starša kregata vsak dan in da to otrok vse sliši ali pa da se ločita, da Lojzka ne bi trpela.

#### 4.2.5 The word *pa* = S/IZPUST/veznik-pa

In these examples the teacher added to the sentence the word *pa*, typically in the role of conjunction. This category currently does not meet the quantitative threshold, but we nonetheless still treat it as a separate category to facilitate analysis and comparison with examples where the teacher removes the word *pa*.

S/IZPUST/veznik-pa:

- V njej zasledimo očeta (Boga) in sinova (ljudi), ki prikazujejo vernike in slabe, ki \_\_ | **pa** se kesajo.
- Ta dva mladeniča \_\_ | **pa** iz Hamleta nista mogla izveliči prav ničesar, zato se je kralju in ostalim še vedno porajalo vprašanje o vzroku.

#### 4.2.6 Other omitted conjunctions = S/IZPUST/veznik-drugo

In these examples the teacher added to the sentence a conjunction or conjunctive phrase (except containing the word *pa*).

S/IZPUST/veznik-drugo:

- Sedaj pa pomislimo, \_\_ | **ali** je res to to, da zaradi ljudi okoli nas, naredimo nekaj čisto nasprotno nam, samo da bi bili ljubljani.
- Vodja tega zavoda Volodja mu začne pojasnjevati kako \_\_ | **da** je on Polski upornik proti Ruskemu vladarju ter kako je po neuspelem uporu prebežal v Slovenijo.
- Antigona naredi samor v svojem sezidanem grobu \_\_ | **in ko** to zve Hajmon se on zabode in umre.

#### 4.2.7 Repeated prepositions = S/IZPUST/predlog-ponovljen

This category contains examples where the teacher adds a preposition to a concatenated string of prepositional phrases. The added preposition can be repeated, or it can be an allophonic variant (*s/z* or *k/h*).

S/IZPUST/predlog-ponovljen:

- Skozi to pisanje sem postala vedno bolj prepričana, da gre tako pri Kajnu in Ablu kot v Priliki o izgubljenem sinu za pravičnost, ne pa \_\_ | **za** usmiljenje.
- Potem se z mamo in \_\_ | **s** sestro odpravim v Slovenj Gradec po trgovinah.

#### 4.2.8 Other omitted prepositions = S/IZPUST/predlog-drugo

This category contains all other examples where the teacher adds a preposition (it also differs from the B/PREDLOG category in that the student in fact did not write a preposition, and only those examples are considered at the vocabulary level if the student does indeed write a preposition and the teacher replaces it with a new one).

S/IZPUST/predlog-drugo:

- Oče ga je sedaj vzel \_\_ | **s** sabo v Ljubljano, kjer ga je vzgajal s trdo roko.
- Postala sem jezna, da sem prišla na to velikanko zaradi trenerja \_\_ | **za** katerega sem mislila, da je smučar.
- Potem pa še: „ \_\_ | **V** To te je prisilila tista smrklija Vlasta, kaj ne?“.

#### 4.2.9 Personal pronoun = S/IZPUST/zaimek-osebni

In these examples the teacher added to the sentence a personal pronoun (perhaps also with a preposition). This category also contains examples with reflexive personal pronouns.

S/IZPUST/zaimek-osebni:

- Zanj je delal, \_\_ | **mu** dajal dajatve —to je bilo žito, opravljal tlako ...
- Zelo si \_\_ | **je** želi in ji reče naj zapeše, ona ga odvrne.
- Horatio želi izvedeti in mu zato priseže, da bo držal besedo \_\_ | **zase**.

Whenever the teacher in addition to a personal pronoun also adds a verb, the example is given two tags.

S/IZPUST/zaimek-osebni in S/IZPUST/glagol-drugo:

- Bal se je svojega očeta, vendar kljub temu ga je imel rad in \_\_ | **ga je** občudoval.

#### 4.2.10 Other omitted pronouns = S/IZPUST/zaimek-drugo

In these examples the teacher added to the sentence a non-personal pronoun. To this group we also ascribe corrections over several words if, in addition to pronouns, they also include prepositions and other words, if they make a phrase with a pronoun at its core.

S/IZPUST/zaimek-drugo:

- V tebi se razraščam, ki govori \_\_ | **o tem** kako se pesnik razrašča v njej, si jo želi.
- Seveda, jaz bi bil bolj če bi bil \_\_ | **pri vsem tem** samo en možen odgovor

#### 4.2.11 Adjective = S/IZPUST/pridevnik

This subcategory is for examples where the teacher added to the sentence an adjective or adjectival phrase.

S/IZPUST/pridevnik:

- Misli je, da je umrl naravne smrti dokler se ni straži prikazal sam duh \_\_ | **pokojnega** kralja.

- Hotel sem prijeli sonce je knjiga, katera govori o Tonetovem \_\_ | **otročkem in mladostniškem** življenju, kako ga je preživel.
- Jesti moramo veliko sadja in zelenjave in paziti moramo na pestrost prehrane in na \_\_ | **pravilno** razmerje posameznih skupin živil in rastlinskih olj.
- Ko sem jaz bila \_\_ | **sama** v taki situaciji, sem storila kot sem zgoraj opisala.
- Ob koncu vojne je ubil Jošta, ker je slednji hotel \_\_ | **Lukeževo** tretino ukradenega denarja.

This category includes omitted added quantifiers that are understood as adjectival in nature.

- menim, da si ti \_\_ | **dve** osebi nista zaslužili biti oče in mati, saj Lojzki nista izkazovala nobene ljubezni.

#### 4.2.12 Adverb = S/IZPUST/prislov

This subcategory is for examples where the teacher added to the sentence an adverb or adverbial phrase.

S/IZPUST/prislov:

- Nekajkrat sem se \_\_ | **zelo** vživel v dogajanje, predvsem z likom Tončkom, saj je bil približno enake starosti in ga razganja tako kot nas.
- Na koncu pa se tragično konča tako, da se je \_\_ | **tragično** utopila v potoku.
- Uprl se je da nebi izgubili oziroma pozabili svoj materni jezik, da bi se \_\_ | **še naprej** ohranjal iz roda v rod.

#### 4.2.13 Particle = S/IZPUST/členek

In these examples the teacher added to the sentence a particle. This category also groups examples where the teacher adds the abbreviation *npr.*

S/IZPUST/členek:

- Tako je bila njena ljubezen nemogoča, saj je Bogomila bila prepričana, da bosta srečna \_\_ | **šele** v onostranstvu.
- Nemški oficir se Berku opraviči za dejanja v vojni, a mu \_\_ | **tudi** pove, da je le izpolnjeval ukaze.
- Tudi kot profesor v šoli se \_\_ | **npr.** zagleda v plašče na obešalih in kar tako ocenjuje kateri je najnežnejši.

#### 4.2.14 Sentence = S/IZPUST/stavek

These are examples where the teacher adds a clause to the student's sentence. Here are grouped examples where the teacher changes a verbless structure into a full sentence by adding a predicate.

S/IZPUST/stavek:

- Literarno obdobje \_\_ | , **v katerega uvrščamo Samorastnike**, je bilo v književnosti med dvema vojnama.

- V odlomku nastopata Simon, ki trdi, da je bil na kliniko sprejet po pomoti in da ni duševni bolnik zraven pa nastopa tudi Volodja, \_\_\_ | **ki je** del zdravstvenega tima ki se zelo rad šali, zlasti z bolniki, včasih malo preveč.

### 4.3 Superfluous constituents

This category contains collected examples where the teacher removed a word, phrase, or longer segment from the sentence. The subcategories are classified by parts of speech, and in some places (in lexical words) also based on content criteria. These criteria were a holdover from the former version of annotation and we have currently kept them under the condition that the subcategories can be easily distinguished from one another.

#### 4.3.1 Literal repetition = S/ODVEČ/ponavljanje

A special subcategory is reserved for examples where the teacher removes a repeated word or phrase, even in instances where the student changed the case in the repetition. These examples are kept separate from other groups due to the stance that the mistakes herein are more closely related to attempts at improving style; furthermore they are, in a formal sense, more easily distinguished than other corrections.

##### S/ODVEČ/ponavljanje:

- Kakor Gregor, pa je bil tudi Berk med vojno na strani partizanov. **Med vojno** | \_\_\_ spozna | Spozna spanskega vojaka Antona, s katerim skupaj preživita hajko.
- Baron ga ne trpi na gradu in ga izžene **iz gradu** | \_\_\_, ampak Tonček na Matičkov nasvet se potuhne in vrne na grad.
- Čez dva tedna grem na morje s starši za štirinajst dni. Letos **na morju** | \_\_\_ upam, da bo zabavno, ko bi vsaj prišel nas presenetiti, to bi bilo zelo zabavno.

In some other examples only part of the phrase is repeated, while another part is only slightly changed or is semantically null. We also include such examples under repetition.

##### S/ODVEČ/ponavljanje:

- Ta odnos je bil nepošten in **s takim odnosom** | \_\_\_ niso ravnali prav, ampak vsaka nezakonska mati je enako vredna kot tista, ki je poročena in ni čisto nič drugačna.
- In mislim, da vsak človek se bi odločil raziskat takšno sporno stvar in poiskušal najti krivca, **za to stvar** | \_\_\_.

Proper nouns, personal pronouns, and auxiliary verbs are a special case, as they appear quite often, wherefore it is more difficult to determine when they are removed to avoid repetition (or more precisely when they were not). These examples are accordingly not annotated as repetitions.

##### S/ODVEČ/samostalnik-lastno-ime:

- Hamletov edini zvesti prijatelj je bil Horacij. **Hamlet** | \_\_\_ se je dobro razumel z Horacijem | **Z njim se je dobro razumel** in je edini prijatelj, ki mu je zaupal vse naloge in se je lahko zanesel na njega.

Similarly we do not include here examples with functional words or parts of speech that define different parts of the sentence, e.g. prepositions in separate phrases.

S/ODVEČ/ponavljanje:

- Ob 18.00 ponavadi pridem domov, kjer še enkrat ponovim snov za šolo ali pa si pripravim večerjo, potem pa gledam televizijo do približno **do** | \_\_ 23.00 ko se odpravim v posteljo.

S/ODVEČ/predlog:

- Črtomir je bojevit, kasneje se počuti krivega za smrt **za** | \_\_ kar zelo obžaluje.

#### 4.3.2 Noun: common noun = S/ODVEČ/samostalnik-občno-ime

In these examples the teacher has removed a superfluous noun (possibly together with a preposition) or a phrase with a noun at its core.

S/ODVEČ/samostalnik-občno-ime:

- Črtomir se zelo spremeni **med prehodom** | \_\_ iz Uvoda v Krst zato, ker po porazu pri Ajdovskemu gradu spozna, da se je utopično boril za samostojnost slovencev, v veri, ampak vseeno mu ostane upanje v Bogomilino ljubezen.

In some cases we must decide whether to tag a correction as rectified repetition or as an ordinary correction to a noun. In the example below the student wrote the same word twice, but the teacher replaced or deleted both. Since neither of the two originally repeated words is left in the text, the first word is annotated as a vocabulary mistake and the second as a superfluous noun.

S/ODVEČ/samostalnik-drugo:

- To knjigo | **delo**, ki jo | **ga** je napisal Sofokles mnogi ljudje večkrat preberejo, saj je **knjiga** | \_\_ izjemno zanimiva | zanimivo in vsebuje vse prvine uspešne zgodbe.

#### 4.3.3 Noun: proper noun = S/ODVEČ/samostalnik-lastno-ime

In these examples the teacher removed from the sentence a proper noun. As mentioned, proper nouns are not annotated as repetitions.

S/ODVEČ/samostalnik-lastno-ime:

- **Špela** | \_\_ ima strogega očeta in mamo. Ko staršev ni doma preživlja čas pri teti Angeli in takrat se tudi počuti bolj varno. Špela ima najboljšo prijateljico Floro, vendar se z njo družijo samo v prostem času.
- Všeč mi je da je **Tone Partljič** | \_\_ preizkusil že veliko poklicev in v njih tudi uspel.

#### 4.3.4 The verb biti = S/ODVEČ/glagol-biti

In these examples the teacher removed from the sentence an instance of the verb *biti*. Though a repeated verb is often at fault here, it is not grouped in the category of repetitions.

#### S/ODVEČ/glagol-bit:

- Prav tako kot antigona sem kršila „zapoved” nekoga drugega in s tem **sem** | \_\_\_ naredila, kar je bilo prav.
- Lojzka **je** | \_\_\_ pa je ležala na blazinastem stolu in se igrala s knjigami.
- Iz vsega tega bi lahko sklepali, da je prijateljstvo s Hamletom zelo težko, a vendarle možno, če bi ta oseba lahko povedala Hamletu svoje mnenje, pri tem pa ne **bi** | \_\_\_ izgubila svoje integritete.

This category has rare examples where the student used a participle of the verb *biti* together with a future-tense form of the verb *biti* and the teacher removed the participle.

#### S/ODVEČ/glagol-bit:

- Torej je pomembno le, da misliš **bil boš** | **boš** to kar si!
- Ta zabava **bo bila** | **bo** nora, pa še srečo imam da bosta starša plačala.

#### 4.3.5 Other superfluous verb = S/ODVEČ/glagol-drugo

In these examples the teacher added to the sentence a verb or verbal phrase.

#### S/ODVEČ/glagol-drugo:

- Luka in Matej sta morala **iti** | \_\_\_ domov saj je bila že pozna ura.
- Menim, da Meursaultu tega ni uspelo, zato je življenje začel jemati kot absurd oz. nesmisel in tako **postane** | \_\_\_ drugačen od drugih.

#### 4.3.6 The word *pa* with another conjunction = S/ODVEČ/veznik-pa-vezniki

These are examples where the teacher removes the word *pa* in concurrence with another conjunction, even if the two words are not located immediately one next to the other.

#### S/ODVEČ/veznik-pa-vezniki:

- Čez mesto teče reka Ljubljanica, ki *pa* je žal zelo onesnažena, vendar **pa** | \_\_\_ društvo Čista Ljubljana, vsako leto organizira čistilno akcijo.
- To pokaže, kako jih ima rad in ta ljubezen ni minila, **pa** | \_\_\_ čeprav ga ni bilo zraven, ko so otroci odraščali
- A ptiček **pa** | \_\_\_ me še kar naprej ”debelo” gleda.

#### 4.3.7 Other examples including the word *pa* = S/ODVEČ/veznik-pa-drugo

This category features all other examples where the teacher removes the word *pa*. most often when functioning as a conjunction.

#### S/ODVEČ/veznik-pa-drugo:

- Župnik **pa** | \_\_\_ je popolnoma drugačen od Jermána.
- Mlajši sin **pa** | \_\_\_ v priliki je zelo nepremišljen, brezkrben in tudi naiven.
- Nato **pa** | \_\_\_ se je **pa** | \_\_\_ vse začelo dogajati hitro.
- Odločil sem **pa** | \_\_\_ se, da bom pokopal Polinejka in to zelo slovesno.

#### 4.3.8 Conjunction at the beginning of a sentence = S/ODVEČ/veznik-začetek

In these examples the teacher removed a conjunction at the beginning of a sentence. This is a matter of cohesion and coherence at the meta-sentence level (linking sentences among each other), where the student begins the sentence with a one-word or several-word conjunction, which the teacher corrects.

##### S/ODVEČ/veznik-začetek:

- S tem se je Linhart norčeval iz njega. **In** | \_\_ ta čas ko se je baron ukvarjal z ženo in Nežko, je imel Matiček še en problem.

Whenever the corrected conjunction is at the beginning of the sentence and is also part of a conjunction string, it is annotated based on whether the teacher's correction removes the whole string or just a part.

##### S/ODVEČ/veznik-začetek:

- **Tako, da** | \_\_ s tem dejanjem ko Lukež ostane pri Kalanu, nam Tavčar pokaže to, kako pomembno je včasih imeti samo prijatelja, tudi če ti ta ne vedno govori le resnice in s tabo ne ravna v vseh primerih pravično.

##### S/ODVEČ/veznik-dvojni:

- **Zato** | \_\_, da oznanja resnico in pravico je tudi bil poslan na svet, saj je pričeval o resnici in pravičnosti.

#### 4.3.9 Doubled conjunction = S/ODVEČ/veznik-dvojni

In these examples the teacher removed a conjunction in a conjunction string (e.g. *zato ker*, *tako kot*, *zato ker*, *kljub temu pa*, etc.).

##### S/ODVEČ/veznik-dvojni:

- V njem pa je napisal sporoča, **češ** | \_\_ da se Baronica z nekom sestaja, da bi tako zbudili njegovo ljubosumje.
- **Zato**, | \_\_ da oznanja resnico in pravico je tudi bil poslan na svet, saj je pričeval o resnici in pravičnosti.
- Nekega poletnega dne ko smo z jadrnico pluli **in** | \_\_ nato smo se ustavili v nekem zalivu.

#### 4.3.10 Other superfluous conjunction = S/ODVEČ/veznik-drugo

Here go all other examples where the teacher removes a conjunction from a sentence. There are many examples where the teacher removes the conjunctions *in*, *ter*, or *kot*.

##### S/ODVEČ/veznik-drugo:

- Gospodična, ki je imela dolge **in** | \_\_ ravne svetle lase in je bila blede polti se mi je predstavila.
- Zato sem si uredil **kot** | \_\_ majhno počitniško hišico.

#### 4.3.11 Preposition = S/ODVEČ/predlog

These examples show where teachers have corrected a superfluous preposition.

S/ODVEČ/predlog:

- Ismena se je zavedala posledic, zato ni hotela z njo. Antigona je bila jezna **na** | \_\_ nanjo, a vendar je to ni ustavilo, saj je svojega brata vendarle pokopala.
- Črtomir je bojevit, kasneje se počuti krivega za smrt **za** | \_\_ kar zelo obžaluje.

Whenever the correction to a superfluous preposition segmented together with another (usually lexical) word, the decision is made based on context. Whenever the preposition is superfluous because a different word was deleted, it is not annotated as a corrected preposition.

S/ODVEČ/zaimek-kazalni:

- Lahko pa le povemo svoja mnenja in moje mnenje **k temu** | \_\_ je da je umor tako hud zločin, da ga ni mogoče izbrisati in na noben način opravičiti, pa čeprav imamo ljudje za seboj toliko slabih izkušenj, da bi nas obvarovale pred krivdo.
- **S tem** | \_\_ ko delaš kar želiš in s tem škodiš marsikdaj drugim, si svoboden, ampak kratiš svobodo drugega.

Whenever the preposition is superfluous, even if another word was not deleted, it is annotated as a superfluous preposition as well as a superfluous lexical word.

S/ODVEČ/predlog + S/ODVEČ/pridevnik:

- **Za napravljeno dejanje** | \_\_ | **Dejanje** po eni strani obžaluje in se kesa, po drugi strani pa ne.

#### 4.3.12 Personal pronoun = S/ODVEČ/zaimek-osebni

In these examples the teacher removed from the sentence a personal or reflexive personal pronoun.

S/ODVEČ/zaimek-osebni:

- Antigona naredi samor v svojem sezidanem grobu in ko to zve Hajmon se **on** | \_\_ zabode in umre.
- Julija mu izpove ljubezen **do njega** | \_\_.
- Ko za Hajmonovo smrt izve njegova mati si še ona sodi sama **sebi** | \_\_.

#### 4.3.13 Demonstrative pronoun = S/ODVEČ/zaimek-kazalni

In these examples the teacher removed from the sentence a demonstrative pronoun.

S/ODVEČ/zaimek-kazalni:

- Njene pozitivne strani so **to** | \_\_, da lahko spomočjo nje dobino iformacije, vsakdanje novice..
- Bronjo ni ustavilo **to** | \_\_, da je Djuro reven in samo avtomehanik.



Whenever a demonstrative pronoun arises in a several-word phrase containing a preposition, conjunction, etc., it is annotated only as a superfluous demonstrative pronoun.

- Mati je otroka nagovarjala **k temu** | \_\_, da kljub vsemu, da ga nihče nemara ga bo ona vseeno ljubila in skrbela zanj.
- Sama meni, da si to zasluži čisto vsak, predvsem pa bi rada poskrbela **za to** | \_\_ da bo njen brat počival v miru in bo tako tudi ona imela čisto vest in bo v miru živela.
- Že tretjič v tem dnevu je naletel **s tem ko** | \_\_ je osvajal svojo ženo, misleč da je Nežka.

#### 4.3.14 Possessive pronoun = S/ODVEČ/zaimek-svojilni

In these examples students have used a superfluous possessive and reflexive possessive pronoun.

S/ODVEČ/zaimek-svojilni:

- To sporočilo pa sem si tudi sama vtisnila v **svoj** | \_\_ spomin, saj vem, da mi bo marsikdaj prišel zelo prav.
- S tem dejanjem sem se prilagodila željam staršev, kajti vedela sem, da jim glede **njihovega** | \_\_ denarja ni popolnoma vseeno.

#### 4.3.15 Other superfluous pronouns = S/ODVEČ/zaimek-drugo

These examples show where teachers have corrected superfluous pronouns, but they fall short numerically of the 35-instance threshold. Often the corrections here target full-form and indefinite pronouns.

S/ODVEČ/zaimek-drugo:

- Tudi jaz sem **istega** | \_\_ mnenja, da za ljubezen ni nikoli prepozno.
- Naslovno misel "Mislim, torej sem!", si razlagam kot **nek** | \_\_ nasvet.
- **Vsi** | \_\_ veseli smo stekli, k njim saj smo jih zelo pogrešali.

#### 4.3.16 Adjective = S/ODVEČ/pridevnik

In these examples the teacher has removed a superfluous adjective (possibly together with a conjunction) or a phrase with an adjective at its core. This category also includes superfluous quantifiers that are understood as adjectival in nature.

S/ODVEČ/pridevnik:

- Odlomek uvrščamo v **ciklični** | \_\_ in impresionistični roman.
- Na Klancu je Živela je mati z **2** | \_\_ hčerkama Nežko in Francko.
- Čeprav nas je malo v primerjavi z drugimi narodi in državami, bi lahko slovenski jezik bil bolj spoštovan **in govoreč** | \_\_.

Whenever a superfluous adjective is segmented together with a superfluous adverb, both superfluous words are annotated.

S/ODVEČ/pridevnik + S/ODVEČ/prislov:

- Kot partizan **najbolj delujoč** | \_\_ v duhu kolektivizma.

#### 4.3.17 Adverb of degree = S/ODVEČ/prislov-mera

In these examples the teacher removed an adverb that intensifies, mitigates, or otherwise expresses degree or extent. The most common here are *zelo*, *malo*, *dokaj*, *veliko*, and their comparative and superlative forms.

S/ODVEČ/prislov-mera:

- Iz odlomka ugotovimo da so bile nezakonske matere v preteklem času **zelo** | \_\_ šibki člen v družbi.
- **Najbolj** | \_\_ glavno vodilo v življenju pa ji je, da bi rada poskrbela, da njeni otroci in vnuki nebi delali istih napak.
- Bog se je zmenil samo za Ablovo daritev, zato je bil Kajen **nekoliko** | \_\_ ljubosumen.
- Hamletovo **relativno** | \_\_ mirno življenje prekine očetova smrt.

#### 4.3.18 Other superfluous adverbs = S/ODVEČ/prislov-drugo

Here we group other examples where the teacher removed an adverb (or adverbial phrase); adverbs denoting locative and temporal relationships are common here.

S/ODVEČ/prislov-drugo:

- Ta usodna ljubezen je ljubezen na prvi pogled in že se Romeo in Julija zaljubita, tvegata vse samo, da bi bila skupaj, **tu** | \_\_ na koncu umreta za ljubezen.
- Nezakonsko mati, ki jo je zapustil oče otroka, ljudje **danes** | \_\_ vidijo v povsem drugačni luči.
- Ko so na se **prav tako** | \_\_ tudi Zmešnjava in Budalo pogovarjata.

#### 4.3.19 Particle = S/ODVEČ/členek

These examples show where teachers have corrected a superfluous (potentially multi-word) particle.

S/ODVEČ/členek:

- Da je ravnal napačno je ugotovil **že** | \_\_ prepozno, ko je **že** | \_\_ bil brez vsega.
- Meni se zdi, da je ta ljubezen žalostna, ker na koncu **tudi** | \_\_ oba umreta.
- **Tako rekoč** | \_\_ mu je bilo žal, da je naredil tako veliko napako in se je kesal svojih grehov.

#### 4.3.20 Clause = S/ODVEČ/stavek

In these examples the teacher removed an entire clause. In some examples a superfluous metatext or ancillary clause was targeted. Mistakes concerning entire clauses are not annotated as mistakes with superfluous content.

S/ODVEČ/stavek:

- Kreona je vprašala, če si se želi kaj več kot le njeno smrt, kar mi je bilo malo smešno. Kreon ji je odgovoril, da mu njena smrt zatošča zato kar je storila **in si ne želi nič več s čemer se ta odlomek pogovora tudi konča** | \_\_.
- Tisti čas smo se otroci veselili in se zabavali dokler mi ni na glavo padel sneg s smreke, ki jo je moj brat potresel, **da je padel name** | \_\_.

#### 4.3.21 Sentence = S/ODVEČ/poved

In these examples the teacher removed an entire sentence. This category contains examples where an entire sentence, from the first capitalized letter to the final punctuation mark.

S/ODVEČ/poved:

- Za domače branje sem si izbral knjigo z naslovom v Sedemnajstem. **O njej moram napisati spis** | \_\_ **no bom, kar začel** | \_\_. Knjiga govori o sedemnajst letni Špeli Kalan.
- In Capulet mu odgovori, da je še mlada ter da je za njega vse. **In kar se bo ona odločila bo to za njega sveto.** | \_\_ je prišel s prošnjo, da prosi Capuleta za Julijino roko.

#### 4.4 Structure

To this group we ascribe examples of substitutions at the structural level (a teacher replaces an individual constituent with another, thereby affecting the syntax of the whole sentence). This category features examples concerning the use of punctuation as related to conjunctions, phrasal syntax, converting phrases to clauses or vice versa, and reworking clauses themselves.

##### 4.4.1 Possessives with od = S/STR/svojina-od

These are examples where the teacher corrects the use of possessive structures built around the preposition *od* such that *od* is deleted or replaced with a different structure.

S/STR/svojina-od:

- Prešernov pogled na usodo zakonske matere je zelo občutljiv, saj je njegova žena Ana Jelovšek dala v rejo tretjega otroka **od njiju** | \_\_, čeprav jo je prosil, naj ga sama vzgaja.
- Bog se je ozrl samo na **daritev od Abla** | **Abelovo daritev**, zato ga Kajn v jezi ubije.
- Hamlet je za smrt svojega očeta odkril morilca Klavdija, za katerega mu je povedal duh **od** | \_\_ njegovega pokojnega očeta in takrat se je odločil, da bo ubil Klavdija.

##### 4.4.2 Possessives with the genitive = S/STR/svojina-rodilnik

These examples show where teachers have corrected a possessive structure built around a noun in the genitive. In the correction the noun is replaced with an agreeing adjectival premodifier.

S/STR/svojina-rodilnik:

- Ko pride v sobo, kjer je ležalo **truplo matere | materino truplo**, obkroženo z njenimi prijatelji, se mu zdi, da ga vsi gledajo in se posvečajo le njemu.
- Hamletu pa to ni bilo všeč saj je sumil Klavdija za **umor očeta | očetov umor**.

#### 4.4.3 Substitution punctuation-conjunction = S/STR/ločilo-veznik

In these examples the teacher replaced a punctuation mark with a conjunction or vice versa. To this category belong examples where the substitution does not affect the number of clauses in the sentence, wherefore it is not considered merging or separating clauses/sentences. Mostly this is the substitution of commas and conjunctions, though examples of switched hyphens/dashes and conjunctions also go here.

##### S/STR/ločilo-veznik:

- Izidor je bil kljub čudnim razmeram v družini, **| in kljub** vsemu nasilju, ki ga je videl in preživel zelo miren in ne grob človek, do očeta ni kuhal nikakršne zamere, kljubtemu, da mu je odsekal prs.
- Ena izmed zelo pomembnih, **| in** zadnje časa vedno bolj cenjenih prednosti je neposredni stiki z naravo.
- Vendar pa je v besedilu razum na zadnjem mestu in zato vsi liki prvo vodijo v svoj **čustveno-duševni | čustveni in duševni** propad nato pa v smrt .

Sometimes due to incorrect segmentation the conjunction appears along with another word in the correction. For instance in the example below we annotated the substitution of comma for conjunction, as well as the omission of a verb.

##### S/IZP/glagol-bitu + S/STR/vejica-veznik:

- Vlasta Korošec ima kratke lase, rjave oči, **| in je** srednje rasti.

#### 4.4.4 Merged clauses = S/STR/združevanje-stavkov

In these examples the teacher has merged several clauses or sentences into one sentence. Such corrections are annotated as “omitted” or “deleted” words/phrases/structures as well as on the punctuation.

##### S/STR/združevanje-stavkov:

- Herod je nekega dne priredil slavje. |, \_\_ **| ki** Salomi slavje ni bilo preveč po godu.
- Sledila je ura športne vzgoje. Tako kot vedno smo se najprej preoblekli. **| in** Nato smo se pogovarjali dokler ni prišel učitelj.
- Pouk se nadaljuje ob desetih. | \_\_ **Pouk se | in** konča pet do pol enih.
- Hamlet se sooča z osebnimi težavami. |, **Povzročila jih je | ki jih je povzročila** smrt njegovega očeta.

Whenever due to the merging of sentences there occurs in the correction a substitution at the vocabulary level, it is not annotated, as it arose due to the changed structure.

##### S/STR/združevanje-stavkov:

- A kmalu mi ga je moja sestra vzela in se je šla kopati v kad, bila sem zelo jezna. **|, Ker | saj** potem še dolgo nisem dobila družga.
- Prebrala sem knjigo MRK. **|, Knjiga | ki** mi je bila zelo všeč, saj je hkrati ljubezenska in napeta.

To this category are also ascribed examples where the teacher merged clauses within one sentence. Here we must be consider changes to the number of clauses. Whenever clauses are merged (at least one of them is completely deleted) and the teacher furthermore reworks one of the clauses, only the merged sentence is annotated (and not the reworked clause), from the stance that the clause was reworked because of the merging.

#### S/STR/združevanje-stavkov:

- Imam nekaj izkušen iz svojega življenja o prijateljici, ki je imela | Tudi moja prijateljica je imela prav tako neurejeno družino.
- Menim, da je tudi pravičen, saj posamezniku dovoli svobodo \_\_ **|, in jim dovoli, da odnehajo | da ostane ali odneha.**
- Odisej in njegovi možje so pa, ko je kiklop zaspal vzeli razžarjeno deblo in mu zapičili globoko v oko **tako, da se mu je kr zasmudila od vročine | in ga oslepili.**

On the other hand we do not annotate the merging of clauses when the clause is simply reworked.

#### S/STR/preoblikovanje-stavka:

- Odisej je želel slišati petje siren zato, ker je slišal, da je njihovo petje pogubilo veliko mornarjev, **zaradi petja, ki jih je | ker jih je petje** vabilo.

#### 4.4.5 Separation of clauses/sentences = S/STR/deljenje-stavkov

In these examples the teacher has broken up one or more sentences or clauses into more than one sentence or clause. We tag the mistake in the punctuation and on the word or phrase that was deleted or added. In doing so, other changes, e.g. at the level of replaced vocabulary, such as in the last of the examples below, are not annotated additionally.

#### S/STR/deljenje-stavkov:

- Še posebej se to pojavlja v priliki o izgubljenem sinu, kjer oče uporablja naraščanje, ko se vrne domov pa tudi sama prilika je po svoje zanimivo literarno besedilo, **|. ki je bilo značilno | Gre** za Jezusove govore za katere je bilo značilno, da so vsebovali primere, nin sicer zato da si ljudje lažje predstavljali „stvari“, ki jih je Jezus razlagal.
- Odgovorila mu je, da je to po božjih navodilih da ima vsak umrli pravico do večnega počitka v posmrtnem življenju, **in da je to storila |. Storila pa je to tudi** iz ljubezni do njenega brata Polinejka.
- Pusti ga, da moli in se skesa, **|. da | Tako** bi bil brez greha in tako šel v nebesa, medtem ko, njegov oče ni šel v nebesa, saj se pred smrtjo ni mogel skesati.

#### 4.4.6 Word/phrase instead of clause and vice versa = S/STR/besedna-zveza-stavek

Here are examples where students write a word or phrase from which the teacher makes a clause. Also here are ascribed examples of the opposite, when the teacher reduces a clause to a word or phrase.

##### S/STR/besedna-zveza-stavek:

- Ko je za to izvedel Kreon vladar, ki je **zelo brezpameten | ravnal zelo brezpametno** saj mu je na koncu bilo žal za svoje dejanje je se začel z antigono pripraviti.
- ... kaj bo naredil z baronom in ko mu je Nežka očitala ji je vrnil **s tem, ko je rekel | z besedami**, da če bi barona slekli in mu vzeli vse premoženje, bi bil manj kot on sam.

#### 4.4.7 Reworked clause = S/STR/preoblikovanje-stavka

In the examples in this category the teacher has reworked a clause into a different clause, often trying to express the same semantic in a grammatically or stylistically more appropriate form (more precise or more concise). In this category “clause” means in the majority of instances a combination of a predicate and at least one other clausal constituent. Reworking that concerns merely a predicate is treated at one of the levels above.

To this category are also ascribed improperly segmented example that cannot be resolved with merely one tag.

##### S/STR/preoblikovanje-stavka:

- **Doleti ga tudi nesreča | Še bolj je nesrečen**, ko svojo zgodbo objavi v neki knjigi, za katero v njihovem kraju ni pričakovati, da bi bila brana.
- Župnik pozna njegove šibke točke, zato tudi rad Jermana zmede, da odreagira z burno reakcijo | zaradi česar se Jerman burno odzove.
- Minilo je nekaj časa, ko je sel znova pritekkel in rekel, da je Antigona bila tista, ki ga je pokopala, saj so jo **videli, ko ga je poiskovala že drugič pokogapati | ujeli pri dejanju**.

If the correction converts a verb to a non-verbal structure and vice versa, it is annotated as S/STR/besedna\_zveza\_stavek. If a verb in the infinitive is substituted by a nominal structure, it is annotated as B/MEN/drugo.

##### S/STR/besedna-zveza-stavek:

- Delo je lirika saj ima verze | lirika v verzih, ima kitice in vsebuje rimo.

##### B/MEN/drugo:

- Zato je še bolj izgubil voljo **živeti | do življenja**.

## 4.5 Additional Annotation

The additional syntactical tags used were taken from the previous annotation system. We retained them where it seemed appropriate, though only as secondary or additional tags. Such annotation provides for the possibility of a more thorough interpretation of correction, especially as regards predicting the rationale behind a correction if it is insufficiently clear from the example itself. Another reason for keeping such tags is that syntactical corrections target style much more than corrections at lower levels. Annotation only at the level of structure further produces certain heterogeneous groups and additional tags thereto can make for more precise and more efficient search operations. It should, however, be emphasized that additional tags are not currently attributed to all data strings, that such tags are limited to those corrections where the teacher removes part of a clause, and that there is some overlap in the usage of such tags. Further work should be devoted to developing these categories more clearly.

### 4.5.1 Pleonasm = S/DOD/pleonazem

This additional tag is used to annotate corrections to words or (partial) phrases containing an element that is redundant or synonymous to another element in said phrase. It is characteristic of such examples that their semantic redundancy is clear even in the absence of further context (e.g. *vrniti se nazaj, zgledati po videzu*). To this category also belong examples of doubling items from the same word family and expressing semantically similar content in different words.

S/DOD/pleonazem:

- Ko se je vrnila **nazaj** | \_\_ je mati ni lepo sprejela, njeno življenje je bilo tudi naprej pusto, ko se je poročila.
- Stopila sem **ven** | \_\_ iz razreda.
- Herod se za njeno bežanje od njega ne zmeni in vztraja **še naprej** | \_\_
- Začeli so se prepiri **in vsa mogoča kreganja** | \_\_ celo pretepi.
- Zagledal sem Antigono obešeno **ter mrtvo** | \_\_.
- Na sprehodu pa sn se spraševal kakšna zgledata **po videzu** | \_\_.
- Matiček se ženi je satirična komedija, kar pomeni, da je besedilo, ki **na smešen način** | \_\_ zasmehuje napake drugih.

### 4.5.2 Superfluous content = S/DOD/vsebina-drugo

In these examples the teacher removed superfluous or repetitive information. Such corrections are often triggered by context in a previous sentence. The category of Superfluous Content also contains examples where a superfluous multi-word phrase paraphrases another word or phrase in close proximity (but does not directly semantically reduplicate it as in the category Pleonasm).

S/DOD/vsebina-drugo:

- Draga Vlasta! **Živjo Vlasta** | \_\_ tako sem bila vesela ko sem te našla ...

- Petruccio pripoveduje Sem Petruccio, in moje življenje se je močno spremenilo zaradi **neke ženske** | \_\_ Katarine.
- Le nekateri ljudje, ki so res močno zaljubljeni bi za svojo ljubezen bili pripravljene spremeniti vero, a če je ljubezen res lahko tako močna **eden do drugega** | \_\_ bi storili tudi to.

In cases of uncertainty to ascribe the correction to Pleonasm or Superfluous Content, the more general one is preferred. This category in the example below is Superfluous Content.

S/DOD/vsebina-drugo:

- Obema družba ni dovolila, da bi bila kar sta saj sta **v sebi** | \_\_ skrivala tisto pravo jaz.
- Moja čustva do tebe so zelo globoka in nislim, da sem zaljubljena u tebe, saj neprestano nislim nate, vedno se ni u glavi prikažuješ ter se te nenoren naveličat gledat, saj si zelo dober **po osebnosti** | \_\_ si ljubeč, postavem ter me vedno spraviš v dobro voljo, čeprav ti tega ne veš in bi ti use to rada povedala u tem pismu.
- Vsi smo posedli **na stole** | \_\_ in razredničarka je spregovorila.

#### 4.5.3 Erroneous content = S/DOD/vsebina- napake

This additional tag is used to define examples where the teacher identified an error in content that the student should fix on their own. Such examples are recognizable in that a word or phrase is deleted, but the correct answer is not written.

S/DOD/vsebina-napake:

- Odlomek spada **med povest** | \_\_.
- Tudi **stražarja** | \_\_ sta na Klavdijevi strani.

#### 4.5.4 Semantically null = S/DOD/pomensko-prazni

In these examples students have used words and phrases that are generally assumed to be problematic for having filler, deictic and discursive elements that are usually from colloquial speech or other (according to the teacher's assessment inappropriate) genre, metatextual information, etc.

S/DOD/pomensko-prazni:

- Zato se je odločil, da se bo pretvarjal, da je zblaznel, saj **mislim, da** | \_\_ se mu je to takrat zdela najboljša rešitev, saj ni hotel, da bi ljudje ugotovili, kaj se njemu plete po glavi.
- **In res** | \_\_ kar je napovedala Medeja, se je tudi zgodilo
- Strinjam se s tem da bi takega vrgli v dosmrtno ječo **ali kaj podobnega** | \_\_.

In some of the examples the segmentation makes the correction look like a substitution, while in reality it is an omitted word or phrase. The example below shows an omitted adverb and demonstrative pronoun, both of which are interpreted as semantically null words.

S/DOD/pomensko-prazni:

- **Tu je pomembno to** | **Pomembno je**, kako so kmetje vztrajali pri maternem jeziku.



## 5 ORTHOGRAPHY

Annotations at the orthographical level primarily target capitalization and words written erroneously either as together or separately. In this category the corpus is also annotated for corrections to punctuation, with the placement of commas contributing to many of the examples. The guidelines do not cover this group.

### 5.1 Capital/lowercase letters

To this category are ascribed examples where the teacher corrects a capitalized initial letter to a lowercase one or vice versa.

#### 5.1.1 Adjectives ending in -ski = Z/MV/pridevnik-ski

In this group are collected corrections to adjectives ending in *-ski* that the student has written with a capital letter.

Z/MV/pridevnik-ski:

- To opeva Tebanski | tebanski zbor.
- Vsi so mi govorili, da so gimnazije težke, še posebej **Škofijska** | **škofijska**, saj je tam tudi nov jezik latinščina.
- Razpravo zoper Matička vodi advokat Zmešnjava, poteka pa v **Slovenskem** | **slovenskem** in ne v **Nemškem** | **nemškem** jeziku, kar je tudi prav.

#### 5.1.2 Other adjectives = Z/MV/pridevnik-drugo

Other examples of corrections to capital letters in adjectives often arise in possessive and in certain other descriptive adjectives.

Z/MV/pridevniki-drugo:

- V **Renesančni** | **renesančni** komediji je veliko aktualnih tem.
- Drama z naslovom Veliki **Briljantni** | **briljantni** valček spada v dramatiko.
- Mislim, da si je Kajn želel predvsem **božje** | **Božje** ljubezni in pozornosti.
- In ravno ta **cankarjev** | **Cankarjev** roman govori o " navidez " vzgledni družini.

#### 5.1.3 Common noun = Z/MV/občna-imena

In this group are collected corrections to common nouns that the student has written with a capital letter and that the teacher has corrected to lowercase. These include examples of writing holidays and similar other typical linguistic pitfalls.

Z/MV/občna-imena:

- Znano je, da se ga je **Baron** | **baron** Naletel skušal znebiti, saj je osvajal Nežko.
- Kocbekova Črna **Orhideja** | **orhideja** je novela stanja, notranjega dogajanja, z redkimi dogodki.
- Za **Božič** | **božič** so kdaj starši prišli po svoje otroke, da bi jih vsaj za kratek čas odpeljali domov.

#### 5.1.4 Personal name with lowercase letter = Z/MV/zemljepisna-imena

In these examples the teacher has corrected the orthography of a personal name that the student wrote with a lowercase letter.

##### Z/MV/osebna-imena:

- Zaradi čiste in iskrene ljubezni, ki jo **črtomir | Črtomir** goji do Bogomile.
- Volodja se izživlja nad **doktorjem | Doktorjem** z besedami in mu pove tudi, da je on odrezal nogo Simonu.
- Bogomila – njegova večna in preljuba ljubezen se je zaobljubila **bogū | Bogu**, zato se je žalosten, z ranjenim srcem odločil za to pot.

#### 5.1.5 Nationality with lowercase letter = Z/MV/narodnost

In these examples the teacher has corrected the orthography of a noun that describes nationality or similar regional affiliation into a capital letter.

##### Z/MV/narodnost:

- Komedija je nastala v času reformacije, ki je za **slovence | Slovence** pomenila pravi preporod.
- Včeraj so se zbrali okoli 7. h zjutraj vsi **sevničani | Sevničani**.
- To, da so bili **nemci | Nemci**, se je zelo videlo v odnosu **slovencev | Slovencev** do družine.

#### 5.1.6 Geographical name with lowercase letter = Z/MV/zemljepisna-imena

In these examples the teacher has corrected the orthography of a geographical name that the student wrote with a lowercase letter.

##### Z/MV/zemljepisna-imena:

- Na **slovenskem | Slovenskem** se je razsvetljenstvo začelo z izdajo Kranjske gramatike, ter končalo s Zoisovo smrtjo.
- Kajn in Abel sta bila sinova Adama in Eve, ki sta bila prva človeka na **zemlji | Zemlji**.
- Vsako leto januarja v **podpeci | Podpeci** postavljajo snežne gradove.

#### 5.1.7 Proper nouns with lowercase letter = Z/MV/stvarna-imena

In these examples the teacher has corrected the orthography of a proper noun that the student wrote with a lowercase letter.

##### Z/MV/stvarna-imena:

- Glavni lik Kocbekove novele **črna | Črna** orhideja je Gregor.
- V **uvodu | Uvodu** je Črtomir optimističen, gleda na svet skozi oči otroka.
- Geodetsko središče mesta je Prešernov trg in **tromostovje | Tromostovje**.

### 5.1.8 Direct speech = Z/MV/premi-govor

Due to their frequency, we have made a special category for corrections to capital and lowercase letters in the use of direct speech. Here are also grouped issues concerning quotes.

#### Z/MV/premi-govor:

- Takrat se mi zazdi, kot da bralec oz. gledalec v sebi razmišlja: „ **mu | Mu** bodo dovolili?”
- Šel sem do mami in jo vprašal: kdaj dobim novega kužka in zakaj? | „ Kdaj dobim novega kužka in zakaj? ”
- „Vlasta, vse nam povej.” **So |,**” **so** se oglasili njeni sošolci s skupnim glasom.
- S svojo dramo Hamlet in slavni izrekom „ **biti | Biti** ali ne biti, to je sedaj vprašanje ”, je sprožil veliko filozofskih vprašanj in razprav.

### 5.1.9 Sentence initial = Z/MV/začetek-povedi

Here we group all other examples where the corrections were at the beginning of the sentence or clause. If the correction targets a lowercase letter as the result of another correction in which the teacher merged or separated sentences, the corrected letter is grouped among related corrections (Chapter).

#### Z/MV/začetek-povedi:

- **to | To** je bil najhujši spor, kar sva jih kdaj imela.
- **ampak | Ampak** nihče ne pride do tako kritične točke, da bi začel pobijati vsevprek.
- Priloge: **dokazilo | Dokazilo** o nalogi.

### 5.1.10 Hypercorrection following a period = Z/MV/hiperkorekcija-ločila

Here are grouped examples where the student wrote a capital letter after a period, but the teacher corrected that to a lowercase letter. Very common are problems with periods in ordinal numbers.

#### Z/MV/hiperkorekcija-ločila:

- 17. **Aprila | aprila** — V soboto je v Sloveniji potekala čistilna akcija.
- 2. **Šolska | šolska** naloga
- Pišemo jo sami, čeprav ... **So | so** stvari, dogodki, ki jih nikoli ne bomo mogli doumeti, in prav je tako.

### 5.1.11 Other problems with initial letters = Z/MV/drugo

To this category are ascribed corrections that fit into no other category. Usually these are examples where the student used a capital letter for a reason that does not accord with typical linguistic pitfalls. Here are also attributed rare examples where the correction is related to T-V distinction.

#### Z/MV/drugo:

- Tu pa **Nastopita | nastopita** Nežka in gospa, ki se zmenita, da bosta zamenjali vlogi.
- Od vse te sreče je to dekle umrlo, srce se je ustavilo od veselja, **Ko | ko** jo je njen Janez zaprosil.
- Zahvaljujem se **vam | Vam** in lep pozdrav.

## 5.2 Together/separate

This category contains collected examples where the teacher corrects an instance of words written together/separately.

### 5.2.1 Verb together = Z/SN/skupaj-glagol

These examples show where teachers have corrected the orthography of a verb, usually to separate it from a written negative particle or a reflexive *se*.

#### Z/SN/skupaj-glagol:

- Ženske tudi pomagajo nekaterim moškim likom, da se **nebi | ne bi** preveč zapletli ali si škodili.
- Ker se torej nova oblast odloči, da je Polinejek izdajalec oz. še bolje, da ga ni, ga prepove pokopati, **sepravi | se pravi**, se ga spominjati in mu izkazovati spoštovanje.
- Očetu namreč niso pomembne posvetne dobrine, ampak ponovno snidenje s sinom, za katerega je **žemislil | že mislil**, da je izgubljen.

### 5.2.2 Preposition together = Z/SN/skupaj-predlog

These examples show where teachers have corrected the orthography of a preposition to separate it from the next word that the student wrote.

#### Z/SN/skupaj-predlog:

- **Nažalost | Na žalost** pa je Kreon preveč zaverovan v svoj prav, da bi razumel njeno mnenje in ljubezen do bližnjega.
- Reševalci so takoj prispeli **napomoč | na pomoč**.
- V delu Krst pri Savici lahko razberemo, **dav | da** v ljubezni lahko nastopi tudi ovira.

### 5.2.3 Preposition separate = Z/SN/narazen-predlog

The issue here is the opposite from the one in the previous category: the teacher corrects the orthography of a preposition written separately.

#### Z/SN/narazen-predlog:

- Takrat **iz za | izza** skale skočijo naši stari znanci XpriimekX.
- **Med tem | Medtem** Tonček skoči skozi okno, ko baron odpre vrata, pa je v sobi Nežka.
- **Na sploh | Nasploh** pa je bil Martin Krpan dober, prijazen, duhovit človek.

#### 5.2.4 Adverb together = Z/SN/skupaj-prislov

These examples show where teachers have corrected the orthography of an adverb, usually to separate it from a written negative particle.

Z/SN/skupaj-prislov:

- Na njo se lahko vedno zanesem, **neglede | ne glede** na to, kaj ali kako hudo je.
- Ljudje so se poskušali **čimhitreje | čim hitreje** asimilirati.
- **Ševedno | še vedno** fizično in grdo rešujejo probleme.

#### 5.2.5 Adverb separate = Z/SN/narazen-prislov

The issue here is the opposite from the one in the previous category: the teacher corrects the orthography of an adverb written separately.

Z/SN/narazen-prislov:

- **Dan danes | Dandanes** je zelo veliko upornikov, saj želijo otroci se pokazati čim starejši v družbi.
- **Vse stransko | Vsestransko** pa mu je bil njegov sin Izidor pravo nasprotje.
- Ta del mi je bil tudi **naj bolj | najbolj** zanimiv.

#### 5.2.6 Adjective separate = Z/SN/narazen-pridevnik

These examples show where teachers have corrected the orthography of an erroneously separated adjective. These are often issued with the prefix *ne-* and the orthography of subordinate compounds.

Z/SN/narazen-pridevnik:

- Poved govori o njuni **ne dopolnjeni | nedopolnjeni** ljubezni.
- Misli, da lahko dobi vse in sam se ima za **več vrednega | večvrednega**, ker je pač baron.
- Nastala je leta 1985 in je prav tako kot Antigona drama z **družbeno kritično | družbenokritično** vsebino.

The issues pertaining to the orthography of coordinate compounds that include hyphen corrections are treated at the syntactical level.

S/STR/ločilo-veznik:

- Vendar pa je v besedilu razum na zadnjem mestu in zato vsi liki prvo vodijo v svoj **čustveno-duševni | čustveni in duševni** propad nato pa v smrt .

#### 5.2.7 Other separate = Z/SN/narazen-drugo

To this category are ascribed all other (less typical and predictable) examples where the teacher corrects separated orthography into a word written together.

Z/SN/narazen-drugo:

- Upam, da bo na dan, **katerega koli | kateregakoli** človeka Bog usmiljen, in bi tako napravil ljudem drugo možnost.
- Vidimo, da se je baronica igrala z baronom in **če prav | čeprav** ni bil nič kriv, se je moral on na koncu zmeraj opravičiti ter osramotiti.
- Tomaž, Denis, Svit so osnovno šolci | osnovnošolci.

### 5.2.8 Other together = Z/SN/skupaj-drugo

To this category are ascribed all other (less typical and predictable) examples where the teacher corrects a word written together into separated orthography.

Z/SN/skupaj-drugo:

- **Karkoli | Kar koli** ga prosim, vedno mi ustreže.
- Cela vas jo ponižuje in želi, da bi otroci, **čene | če ne** drugega, imeli vsaj nekaj od očeta.
- Z neizmerno ljubeznijo se zdaj poslavljam in upam, da mi **natopismo | na to pismo** odpišeš.

### 5.3 Abbreviations = Z/KR/drugo

To this category are ascribed examples where the teacher corrects the orthography of an abbreviation. Here are often issues concerning whether to write a phrase in abbreviated or full form, with rarer corrections to the manner of abbreviation.

Z/KR/drugo:

- Da nismo preveč osamljeni, si nabavimo kakšno domačo žival, **npr. | na primer** papagaja.
- Pozneje v 21 **sto. | st.** pa je zdaj vsak lastnik televizije.
- Nagrada je bila **cede | CD** za na računalnik.

### 5.4 Numbers = Z/ŠTEV/drugo

To this category are ascribed examples where there are problems concerning the orthography of a number, e.g. deciding between writing a number symbolically or in words, using periods in writing dates or ordinal numbers, the use of hyphens or dashes for declension or to mean *from—to*.

Z/ŠTEV/drugo:

- Na **1. | prvi** šolski dan smo šli v nov kulturni dom gledat film Deklica in lisica.
- Sodeloval je v **30. letni | 30-letni** vojni
- Spada v obdobje romantike, ki traja od leta **1830 —1849 | 1830 do 1849**.

## 5.5 Punctuation

Of all the groups currently (for versions 2.0 and 3.0) addressed in the Šolar corpus, the punctuation group is the only one that has not been fully manually reviewed and categorised. This is due to the enormous number of corrections, mainly related to the placement of the comma, but sometimes also to the selection or the form of other punctuation marks. In the preparation of the Šolar 1.0 corpus, 1,500 sentences with missing or redundant commas were manually reviewed and categorised. These can be found under the tags that include the string *vzorec-vejica* ('*sample-comma*'). Other punctuation errors are currently unclassified. Before continuing with the current annotation system, an analysis of the designed categories should be carried out and, if necessary, the labels should be upgraded. The remainder of Section 5.5 thus mainly explains the labels that were designed in the 1.0 version of the corpus (see the monograph Kosem et al., 2012: 46-47 for more information).<sup>9</sup>

### 5.5.1 Unclassified punctuation corrections = Z/LOČ/nerazvrščeno

The vast majority of the punctuation corrections (more than 13,000) are unclassified. Among them are various cases of misplacing (in particular) a comma, i.e., cases where the teacher has added a missing comma or removed an unnecessary comma. Corrections to the use of other types of punctuation marks are also among the problems, but less frequently.

- A ker se Tonček, | \_\_ barona boji, zbeži skozi okno pred njim.
- Če imaš vse to, te v življenju obdaja veliko ”| „prijatelj” **prijatelj**” in ljudi, ki te ”| „imajo radi”.

### 5.5.2 Comma before subordinate clauses = Z/LOČ/vzorec-vejica-stavki

These are examples where the problem of comma placement is found before the subordinate clause.

- Čakala je na odgovor | , vendar ga dosti časa ni dobila.

### 5.5.3 Comma between parts-of-speech = Z/LOČ/vzorec-vejica-stavčni-členi

These are examples where the problem of comma placement is found between different parts-of-speech.

- Pri svojem začetnem upiranju, | \_\_ je bil Simon povsem nemočen.

---

<sup>9</sup> KOSEM, Iztok, STRITAR KUČUK, Mojca, MOŽE, Sara, ZWITTER VITEZ, Ana, ARHAR HOLDT, Špela, ROZMAN, Tadeja, 2012: *Analiza jezikovnih težav učencev: korpusni pristop*. E-izdaja. Ljubljana: Znanstvena založba Filozofske fakultete. <https://e-knjige.ff.uni-lj.si/znanstvena-zalozba/catalog/view/229/329/5311-1>.

#### 5.5.4 Comma and multi-word conjunctions = Z/LOČ/vzorec-vejica-vezniki

These are examples where the problem of comma placement is linked to the structure of multi-word conjunctions.

- Abel se trudi pridobiti naklonjenost svojega starejšega brata, | \_\_ med tem, ko ga Kajn odkrito sovraži.

#### 5.5.5 Comma and comparative structures = Z/LOČ/vzorec-vejica-kot

This group contains cases where the comma problem occurs in comparatives and adverbial phrases with *kot* ('as').

- Ni ji bil tako predan, | \_\_ kot pa Bogomili, zato je pristal na krst.

#### 5.5.6 Comma and appositions etc. = Z/LOČ/vzorec-vejica-pristavki

This group includes cases where the comma problem occurs in connection with appositions, semi-clauses, *pa*-clauses, or inserted non-verbal clauses.

- Ima dva prijatelja | , Tomaža in Denisa.

#### 5.5.7 Comma and inserted subordinate clauses = Z/LOČ/vzorec-vejica-vrinjen-odvisnik

This group contains cases where the comma problem occurs before or after inserted subordinate clauses.

- Naslov besedila, ki sem ga prebral | , je Princeska z napako.

#### 5.5.8 Comma and coordinate phrases = Z/LOČ/vzorec-vejica-priredja-zvez

This group contains cases where the comma problem occurs in coordinate phrases.

- Danes se pa vsem nekam mudi in dela se vse hitro, | \_\_ ter slabo.

#### 5.5.9 Comma and coordinate clauses = iZ/LOČ/vzorec-vejica-priredja-odvisnikov

This group contains cases where the comma problem occurs before or after coordinate clauses.

- Izžene Kajna in, | \_\_ ko vidi, da se resnično kesa, mu odpusti.

#### 5.5.10 Comma in adjective strings = Z/LOČ/vzorec-vejica-pridevniški-niz

This group contains cases where the comma problem occurs inside adjective strings.

- Imel pa je tudi mlajšega | , pravega brata Benčija, ki ga je večkrat peljal v vrtec.



### 5.5.11 Comma and predicate ellipsis = Z/LOČ/vzorec-vejica-elipsa-povedka

These are examples where the problem of comma placement is found in sentences with predicate ellipsis.

- Opisuje tudi Ožbeja, ki je v precepu, saj ima Meto zelo rad | , hkrati pa tudi očeta in zapuščino.

## 6 RELATED CORRECTIONS

To this category are ascribed all examples are not corrections *sensu stricto*, but where the teacher's change occurs as a result of a correction elsewhere in the text. Here it is important that such encroaches upon the text are classified separately from true corrections to students' linguistic choices, so that the data might be properly annotated for appropriate interpretation. Related corrections come in a multitude of types and previous categorization schemes were accordingly aggregated at various linguistic levels, which in some places remains inconsistently applied. It is our goal that in further work such related corrections would be regularly annotated along with all others, and that the segmentation of actual corrections would be adapted to this accordingly. Common sense is the guiding principle in determining what the fundamental correction is and what the related one is in the current corpus. Below are a few examples for each level.

### 6.1 P/OBL/drugo

The core or primary correction in the examples below is annotated with red, while the related correction is annotated in gray. Related corrections are often the result of a substituted preposition, e.g. in relation to questions of a verb's valence. The core of the correction in these cases is found in the preposition, and the changed categorical characteristics of lexical words are treated as related corrections.

- Hitro so pluli nazaj k | v svoji | svojo domovini | domovino.

This holds true, i.e. that the preposition is the core of the correction, when in correcting a case form the teacher also changes a sentence's constituents and the role of the relationships in a sentence (e.g. from circumstantial or absolute information to the sentence's subject).

- Kot sem že v uvodu napisala, je bil Polikarp trd kot hlod; že v | \_\_ **samem** | **sam Tavčarjevem** | **Tavčarjev opisu** | **lopis**, kako premika smrečje ali kako jezdi vihravega žrebca, nam lahko da misliti, da je res bil trd kot skala in da mu je le malo stvari prišlo do živega.

Other fairly easily discernible examples include the teacher replacing a verb, prompting a change to the categorical characteristics of a nominal phrase.

- Je zelo vraževeren, saj **verjame** | **ga plaši vsaki** | **vsaka besedi** | **beseda** preroka, ki se ga tudi malo boji, čeprav on to zanika.

In examples where the teacher either removes or adds *se/si* and consequently the verb's reflexivity, that correction is considered core:

- Padla je in vsi so **se | \_\_\_ ji | jo** še bolj zasmehovali.

If the teacher adds in a new constituent to the sentence that changes the syntactical structure and thus categorical characteristics, as shown below, such an insertion is considered the core correction.

- Ko je \_\_\_ | **Odisej Polifem | Polifema** oslepel | oslepil, se je močno zameril Pozejdonu.

Corrections at the level of infinitive–finite verbal form are often the result of other syntactical corrections and are thus grouped among related corrections.

- Pa vendar mi je bil ta lik zanimiv, saj ga je Linhart najbrž hotel takega ustvariti, **saj je hotel s to | da bi s** komiko pravi še bolj **poudariti | poudaril**, da je to njegovo delo komedija.

## 6.2 P/SKLA/osebik

A special group is currently reserved for examples where the teacher corrects word order that arose as a result of corrections to a sentence's (un)expressed subject.

- **Vsi | \_\_\_ smo se razveselili | razveselili smo se** in odhiteli po vrečke, da bomo nabirali školke.
- \_\_\_ | Sam Kot otrok sem se | sem se kot otrok rad igral z sestrinimi stvarmi.

## 6.3 P/SKLA/drugo

At the syntactical level there are also other examples of related corrections that are the result of a teacher reworking a phrasal or sentence-level structure. Often this entails corrections to word order that arises due to the substitution of words, e.g. due to a change from a pronoun to a lexical word.

- **Iz izkušenj | \_\_\_ sem navajena | Navajena sem** na to, da naj spis navadno začnem z mislijo, tokrat ga bom začela na malce drugačen način.
- Pridružil se jim | Partizanom je prostovoljno | Partizanom se je pridružil prostovoljno, vendar se ne pusti očarati voditeljem in visokim ciljem.

## 6.4 P/ZAP/mala-velika

At the orthographical level are grouped related corrections that occur due to the merging of sentences and subsequent change to capitalization.

- Poleg zgradbe je tu še bistvena romantična značilnost, ki se kaže v Krstu, kajti Uvod je sam po sebi bolj epski, | . **razkol | Razkol** med idealom in stvarnostjo ter usoda, ki se preigrava z junaki v katerih najdemo obrise pesnikovega čutenja, saj je tudi sam doživel veliko izgubo prijatelja Čopa in ljubljene Julije, kar ga je pahnilo v romantično resignacijo, iz katere si je pomagal predvsem z izlivanjem čustev na papir in tako je prav zaradi trpljenja nastala ena najboljših Prešernovih in slovenskih pesnitev.

## 7 ILLEGIBLE AND DUBIOUS EXAMPLES

### 7.1 Illegible examples = N//necitljivo

Whenever an example features an error or correction annotated as illegible or anonymized proper noun, it is annotated as illegible.

N//nečitljivo:

- V vojni **§§§mor§§§** | **moraš** ubiti človeka v sebi.
- **XImeX** | **ji** so začeli čistiti iz pljuč vodo ter oživljanje.

### 7.2 Dubious examples = N//preveri

Examples to be further verified are annotated as seriously extraordinary in terms of their correction (where it is possible to assume that there was an error in transcription), as well as examples where the correction is the same as the error itself.

N//preveri:

- Tako je **odisel** | **odesj** slišal petje ne da bi preusmeril ladjo. Odisej je ves čas zvit in domiseln s tem mu je ospelo priti do cilja.
- Ta odlomek govori o Metinih in Ožbejovih **otrocih** | **otrocih**.

In annotating we sometimes find a correction that is unjustified, incomplete, or otherwise suspicious in terms of whether the error and correction were appropriately transcribed in the corpus. Currently such corrections are nonetheless considered appropriate as regards annotation, but it is good to tag them separately for verification against the original.

- Jaz sem se mogel začeti **učiti** | **učit** za spis.
- Tako da ni dovolj, da le mi držimo **svoja** | **svojega** prepričanja.

## 8 DECIDING BETWEEN FRINGE CASES

This chapter discusses a few sets of examples where a dilemma arises as to how to group corrections at various levels of annotation.

Examples where the teacher replaces a whole word in their correction are grouped only at the level of vocabulary item. This principle also applies to substitutions of similar words transcending parts of speech, i.e. word families, e.g. adjective and adverb, and changes to whether a word is written together or separately, inasmuch as this changes its part of speech.

B/MEN/besedna-družina:

- Veliko sprememb se zgodi tudi **nepričakovanih** | **nepričakovano**, ki nas lahko zelo razveselijo lahko pa tudi razžalostijo.

Examples where the teacher replaces one word with another are categorized only at the level of vocabulary, even if the substituted word is spelled incorrectly.

B/GLAG/drugo + B/DOD/zaznamovano:

- Ofelijeve smrt je povzročala bolečine za Hamleta, vendar se je najprej **obadal | ukvarjal** s svojim maščevanjem.

If the teacher replaces a word in the correction and at the same time it is unambiguously clear that they also changed its categorical characteristics, such a correction is annotated at the level of vocabulary as well as morphology.

O/KAT/oseba + B/GLAG/drugo:

- Hajmon mi je zagrozil, da če ne **spusti | oprostim** Antigone bom izgubil tudi njega.

The above principle applies only to examples where the replaced word is paradigmatically regular (*spusti – oprostim*). If the word is irregular (e.g. written with non-standard spelling), the goal is to define this correction with the combination of tags that seems the most logical. For instance, the example below tags a combination of spelling issues (*d-n*) as well as the wrong case in the ending.

Č/KONZ/menjava-drugo + O/KAT/sklon-mo:

- Saj pripovedovalec ima svoje mnenje, ki ga spoštujem, ker ni z **dobenem | nobenim** načinom užalil te osebe.

When the correction targets individual letters that can be interpreted as either a morphological or vocabulary issue, the higher-level category is given precedence.

O/PAR/glagolska-osnova:

- V tem dejanju, avtor kritizira in opozarja, da smo se slovinci, bratje, med seboj **pobivali | pobijali** v drugi svetovni vojni.

When it cannot be sufficiently determined whether the error is in spelling or vocabulary, the context and the frequency and similarity of other incorrect or corrected letters is used as a guide. In the example below it seems less likely that the student accidentally chose the wrong word than that they had trouble spelling the right one.

Č/KONZ/menjava-td:

- Kljub temu, da je bila prej skupina ljudi za **trugo | drugo** stranko, se nekatere da prepričati z dobrimi argumenti.

In principle all corrections to endings are categorized as morphological. If the student did not write an ending at all, it is treated as a spelling mistake (omitted vowel, consonant, or syllable):

Č/VOK/izpust:

- Poznamo ljubezen do Očeta **kater | katero** najdemo v Hamletu, vendar je bolj izražana kakor v aktualnem svetu.

Also considered spelling mistakes are those corrections that only superficially seem as if they belong to the ending. The example below features an instance of the final letter of an adverb where it is impossible to unambiguously determine whether the morphological issue is paradigmatic or categorical, though by analogy it could be determined that the reason could lie in the writing of the letter *a* or *o*.

#### Č/VOK/menjava-ao:

- Človek se zna telesno predat, **duševna | duševno** pa ni zmožen vsak.

Often there are dilemmas in corrections to forms that are clearly incorrect but that are hard to confidently place within categorical errors (e.g. choice of gender, case, number, etc.) or at the level of spelling (because the correction is not indicative of a typical spelling mistake). To ensure consistency and to avoid excessive subjectivity in deciding about such examples, these corrections are annotated at the morphological level. “Categorical problems” are determined such that the tag is as elementary as possible: if several combinations of categorical corrections can be ascribed, the simplest one is to be preferred. The example below for instance tags an incorrect case (as using the dative form instead of the nominative), not as an error in gender or number.

#### O/KAT/sklon:

- Ko je dobila 3 otroke, jih je morala **sami | sama** preživljati in vzgajati.

At the level of spelling are grouped examples where the student omits an unaccented *e* in an ending such as *-ega* or *emu*, while committing no other error at the level of category or paradigm.

#### Č/VOK/izpust:

- Glavni junak je fatima živi v zelo slabem okolju, ima velike težave, v soli jo zafrkavajo, nima **nobenga | nobenega**, da bi ji svetoval, pomaga pa ji učiteljica Kodarlajsa.

Problems unrelated to a word’s morphology but to the orthography of an undeclined word or root (unless one of the identified problems with root transformations) as by default treated as spelling mistakes:

#### Č/KONZ/izpust:

- V **niju | njiju** ni imel tako dobrega mnenja kot o Horaciju.

There are also dilemmas in distinguishing the morphological level from that of vocabulary. In such instances the appropriate placement within the system is sought through the question: “where would it be best to find this annotated mistake?” By the same logic the example below is annotated as a problem with a capitalized initial and an atypical non-verbal ending, not as a vocabulary issue, i.e. the substitution of the words *petek* and *petka*.

#### Z/MV/lastno + O/PAR/neglagolska-končnica:

- Kmalu se vrnejo tujo ljudstvo, s **petko | Petkom** v boju z nimi rešiva vjetnike.

The example below is annotated as an error in number and at the vocabulary level (the masculine form of a noun is substituted for its feminine form).

#### O/KAT/spol + B/SAM/občno-besedišče:

- Romeo in Julia govori o tragediji, ki je za to tragedijo bila kriva njuna družina, ker sta bili največji **sovražniki | sovražnici**.