Stanford Typed Dependencies: Slavic Languages Application

Katarzyna Marszałek-Kowalewska[1], Anna Zaretskaya[2], and Milan Souček[3]

[1] Adam Mickiewicz University, Poznań, Poland
k.marszalek.kowalewska@gmail.com
[2] University of Málaga, Spain
azaretskaya@gmail.com
[3] Lionbridge Technologies, Inc., Tampere, Finland
milan.soucek@lionbridge.com

Abstract. The Stanford typed dependency model [7] constitutes a universal schema of grammatical relationships for dependency parsing. However, it was based on English data and did not provide descriptions for grammatical features that are fundamental in other language types. This paper addresses the problem of applying the Stanford typed dependency model for Slavic languages. Language features specific to Slavic languages that are presented and described include ellipsis, different types of predicates, genitive constructions, direct vs. indirect objects, reflexive pronouns and determiners. In order to maintain cross-language consistency we try to avoid major changes in the original Stanford model, and rather devise new applications of the existing relation types.

Keywords: dependency parsing, treebank, Slavic languages, Stanford typed dependencies.

1 Introduction

There has been a major development in building language resources in the last couple of decades. Various annotated language corpora could be collected thanks to newly developed computational linguistics methods applied to large volumes of data available via the Internet. Also, specialized corpora like dependency treebanks are now available for many languages. Monolingual dependency treebanks have been recently built for a majority of Slavic languages as well (e.g. [1] for Russian; [13] for Czech; [12] for Slovenian; [26] for Croatian; and [30] for Polish). One of the next targets for building language resources is preparation of multilingual data with consistent annotation format that could be used for cross-language research and mainly for Natural Language Processing (NLP) tasks. In terms of dependency treebanks, similar tasks have been tackled in the last decade in the series of the Conference on Computational Natural Language Learning (CoNLL) shared tasks, where various approaches to parsing multilingual data from different treebank sources were presented in [4, 14, 17, 24]. Lately, there also have been activities related to building universal

A. Przepiórkowski and M. Ogrodniczuk (Eds.): PolTAL 2014, LNAI 8686, pp. 151–163, 2014.

multilingual treebank resources [16], using unified Part of Speech (POS) [20] and dependency relation (deprel) [7] schemas. For similar cross-language activities, such unified annotation tag sets are required in order to describe universal grammar relations in languages of different types. The Stanford typed dependencies model (SD) [7] was designed to provide a simple description of the grammatical relationships in a sentence that also could easily be used for tasks not strictly related to linguistic research. This dependency model seems to be suitable for multilingual tasks where grammatical relationships in different languages need to be transformed into a universal schema while a certain level of simplification needs to be introduced. The SD model was initially defined for English and built on English reference data, which means that its application to other languages brings challenges in any cases where new types of grammatical relations need to be handled. Some improvements for the SD schema already have been suggested in various works, both by the original authors of the model (see e.g. [8–10]) and by teams that have been using SD when building treebank data (e.g. [16]). Furthermore, SD was developed on the basis of Wall Street Journal data, i.e. newswire data, which may lack constructions that are characteristic for other data types. To be more precise, although the SD model provides a good description of basic grammatical relations such as subject, object, noun phrase relations, or subordinate clauses, it lacks coverage for questions, discourse particles, ellipsis [9], or grammatical relations which are expressed by bound morphemes in morphologically rich languages. In this paper, we present the SD application for Slavic languages as it was defined for Slavic treebanks that have been built for the Universal Dependency Treebank project.[1] Since dependency treebanks following the SD model are already available for multiple languages (see [5] for Chinese; [15] for Finnish; [22] for Persian; [28] for Hebrew; [3] for Italian; [16] for German, English, Swedish, Spanish, French and Korean, and [23] for French, Spanish, German and Brazilian Portuguese), including Slavic languages resources to this group would be beneficial in terms of extending the language coverage for the cross-language studies and multilingual NLP applications. In the following sections we present problematic examples and provide proposals for applying Stanford dependency model on grammatical relations that are not covered in the basic SD model, but need to be taken into account for Slavic languages.

1.1 Slavic Language Peculiarities

Slavic languages constitute one of the major modern language families in the world. It is the fourth largest sub-family within Indo-European languages with around 300 million speakers [25].

Word order in Slavic family is more free than in Germanic languages, i.e. the order of major constituents (subject, object) is determined more by pragmatic rather than syntactic factors. The major factors that influence Slavic word order are conventional word order, constituent structure and the marking of grammatical relations by means other than word order [25]. However, the order within individual constituents is more fixed, e.g. demonstratives and numerals usually precede nouns (but not always).

[1] `https://code.google.com/p/uni-dep-tb/`

All Slavic languages retain a rich set of morphological categories. As Comrie et al. claim [6] Slavic morphology is mainly fusional, i.e. an affix can combine a number of grammatical categories. In contrast to English, genitive case in Slavic languages is expressed in terms of declension. Another typical Slavic feature is the use of impersonal constructions which can cause difficulties in differentiating between subject and direct object.

The following subchapters provide more detailed information on a few specific features of Slavic languages with regard to applying the SD model. We will focus here on determiners, modality, reflexive pronouns, copular verb ellipsis, difficulties in differentiating between subject and direct object, genitive constructions, and indirect objects.

2 Slavic Application

2.1 Determiners

A determiner is a word that modifies a noun or a noun phrase, e.g. *a, the* in English, or *der, die, das* in German. As a POS category, determiners usually include articles and pronouns (demonstrative, interrogative or possessive pronouns) [29].

In the basic SD model the deprel *det* is used to cover the relation between the head of an NP[2] and its determiner [7]:

The man is here	*det*(man, the)
Which book do you prefer?	*det*(book, which)

For Germanic languages the most obvious determiner is the article. Although Slavic languages lack articles (apart from Macedonian and Bulgarian), their sets of determiners are quite extensive due to rich morphology, i.e. declension of determiners is sometimes known as 'special adjective declension' [25]. To illustrate this, let's compare the demonstrative pronoun *this* in English and Polish:[3]

Eng: [*this*]
Pol: [*ten, tego, temu, tym, ta, tej, tę, tą, to*]

Following the basic SD model, for the Slavic languages application, the relation between demonstratives and interrogatives and the noun that they directly modify is expressed as *det*:

Cz: *Znám toho muže* "I know that man"	*det*(muže, toho)
Cz: *Který den je dnes?* "What is the day today?"	*det*(den, který)

[2] NP stands for a noun phrase. This construction has either a noun or a pronoun as its head [29].
[3] We use following language abbreviations in examples: Eng for English, Cz for Czech, Pol for Polish, Rus for Russian and Slk for Slovak.

The SD model also provides a label for the relation between the head of an NP and a word that precedes and modifies the meaning of the NP determiner (*predet* deprel), e.g.

All the boys are here *predet*(boys, all)

Also in Slavic languages the relation *predet* can be used with words like *all*, *whole* etc., if they precede another determiner, e.g.:

Cz: *Všichni naši přátelé* "All our friends"	*predet*(přátelé, všichni)
Rus: *Все эти люди* "All these people"	*predet*(люди, все)
Pol: *Wszyscy ci ludzie* "All those people"	*predet*(ludzie, wszyscy)

The interesting point is that in some cases, the reverse order of determiners is also possible in Slavic languages. For such multiple determiners, we decided to keep *det* assigned to the demonstrative pronoun and *predet* to other pronouns.

Pol: *Ci wszyscy ludzie* "All those people"	*predet*(ludzie, wszyscy)

The Universal Dependency Treebank [16] proposes to merge *det* and *predet* into just one deprel *det*. If such application is preferable, *predet* easily can be converted into *det* for Slavic treebanks as well.

Although possessive pronouns are often considered determiners as well, the basic SD schema uses the deprel *poss* to describe the relationship between possessive pronouns and the noun that they modify, e.g.:.

Pol: *Moja siostra jest tu* "My sister is here"	*poss*(siostra, moja)

2.2 Reflexive Pronoun

A reflexive pronoun refers back to the subject of the clause in which it is used, e.g. *myself* in English, *sich* in German or *si* in Italian [29].

In Slavic languages the reflexive pronoun may take several forms. It can function as an object-like reflexive pronoun or as a purely reflexive marker of the related verb. In East Slavic, the once-present reflexive-as-clitic has disappeared and now it is expressed as a verbal suffix joined to the verb, e.g.:

Rus: *Она одевается* "She dresses herself"

In other languages, the reflexive acts as a clitic that may be found at different positions, e.g.:

Pol: *Musisz się umyć*	"You need to wash yourself"
Pol: *Musisz umyć się*	"You need to wash yourself"
Pol: *Drzwi otworzyły się*	"Door opened (itself)"

The basic SD model does not provide a special deprel representing the relation between a reflexive and its head (verb). For the Slavic application, depending on its syntactic function, the reflexive pronoun is treated either as a functional pronoun with the label *dobj* or *iobj*, when it has a recognizable pronoun function, e.g.:

Pol: *Piotr umył się* "Peter washef himself" *dobj*(umył, się)

Or as a particle with deprel *prt* in cases when it is purely reflexive, e.g.:

Pol: *uśmiechnąć się* "smile (oneself)" *prt*(uśmiechnąć się)

In the latter example, the reflexive cannot be treated on the basis of syntactic functions; it does not behave as an object. Therefore, we decided to use the deprel *prt* which in the original SD model was used to identify phrasal verbs and described the relation between a verb and its particle, e.g.:

They broke up *prt*(broke, up)

2.3 Copular Verbs Ellipsis in Russian

Copular verbs or copulas are verbs with one complement that serve as a link to what the referent of the subject is or becomes. The most common copula in English is *be* ("*The Earth is round*"). Other verbs used as copulas in English provide additional meaning to the mere linking, like the verbs *become, appear, seem, feel,* etc. [11].

The SD manual [7] defines the *cop* relation as 'the relation between the complement of a copular verb and the copular verb'. In this relation, the copular verb depends on its complement:

Bill is big *cop*(big, is)

In the standard version of the Universal Dependency Treebank project [16], some function words, such as copulas and adpositions, are treated as heads of their complements. We follow this same approach for the Slavic SD application. This means that the copula complement depends on the copular verb. In addition, we discarded the *cop* relation as well as deprels *acomp* and *attr*. Instead, we introduce the dependency label *scomp* (subject complement). The new label *scomp* is defined as 'a verb complement that refers to the subject of the clause. If the complement is capable of inflection, it will agree with the subject in number and gender'. Thus, the nature of the verb (i.e. whether it is copulative or not) is not crucial. This allows the verb to occupy its natural position in the root node of the sentence, while both the subject and the complement depend on the verb, e.g.

Bill is big *scomp*(is, big); *nsubj*(is, Bill)

There seem to be no drawbacks in this representation until it is applied to the Russian language, where the main copular verb *быть* "to be" is almost always omitted in present tense. Consider the three examples provided below, all of which contain a copulative structure. The ellipsis occurs in sentence (1), which on a semantic level, has the verb "to be" in present tense. However, the same sentence in past tense (2) does contain a verb form, as well as sentence (3), where another copular verb is used instead of the verb "to be".

(1) Rus: *Этот студент – лингвист* "This student is a linguist"
(2) Rus: *Этот студент был лингвистом* "This student was a linguist"
(3) Rus: *Этот студент работает лингвистом* "This student works as a linguist"

In the case of copula ellipsis, both the subject *nsubj* and subject complement *scomp* are left without a head node, therefore a decision has to be made to define their relationship. One of the two apparent solutions is to consider the subject as the head of an *scomp* relation with the complement: *scomp*(студент, лингвист). The other solution is to have the subject depend on the complement: *nsubj*(лингвист, студент). The two dependency structures are shown in Figures 1 and 2.

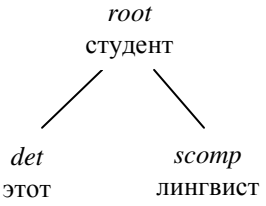

Fig. 1. scomp structure for copula ellipsis

Fig. 2. nsubj structure for copula ellipsis

We concluded that the relation between the verb and complement is closer than the one between verb and subject because – in terms of phrase constituents – they are both parts of a verb phrase and thus form one unit. Therefore it seems more natural, when the verb is omitted, to fill its *root* position by the complement. As a result, we selected the second solution for the Slavic SD application.

2.4 *Nsubj* and *iobj* in Russian

In Stanford typed dependencies, a nominal subject (*nsubj*) is a noun phrase which is the syntactic subject of a clause [7]. Some Russian verbs and constructions require a subject in dative or genitive case. Therefore, it is often difficult to decide whether a noun in oblique case is a subject or an indirect object (*iobj*). Similar examples can be found in Polish as well:

Rus: *Марии нет дома* "Maria(*gen*) is not at home"
Rus: *Ему нельзя выходить* "He(*dat*) is not allowed to go out"
Pol: *Jemu nie wolno wyjść* "He(*dat*) is not allowed to go out"

For the Slavic SD application, we propose to enlarge the definition of *nsubj* so that it also covers the most frequent constructions where the subject is used in oblique case:

- A genitive NP depending on the predicative word *нет*, which in fact is the negated copular verb *to be* in present tense. This also works for all negated structures with verbs expressing existence:

  Rus: *Здесь никого нет* "There is nobody"
  Rus: *Письма не пришло* "The letter didn't come"

- Dative NPs that depend on PRED words or adverbs with predicative function:

  Rus: *Мне нужно уйти* "I need to go"

- Dative NPs with impersonal verbs in third person singular:

  Rus: *Ему пришлось подчиниться* "He had to obey"

In all 3 cases, the non-nominative subject is still analyzed as *nsubj*. In other cases, dative and genitive NPs are treated as regular indirect objects.

2.5 Genitive Constructions

Slavic genitive constructions have been the object of syntax and semantics studies for a long time. Below, we outline the issues that arose in relation to the SD model application to Russian genitives.

For the Slavic SD application, we introduce a genitive modifier relation *gmod* to the SD model as a genitive attribute that modifies an NP. Consider the following:

Rus: *ножка стула* "leg of the chair"
Pol: *pisk opon* "screech of tires"

The problem occurs, however, when a noun in genitive case appears in other contexts that are different from *gmod*. Thus, NPs in genitive case are used in the position of direct object in negative constructions [2]:

Rus: *Я не заметил водки на столе* "I didn't notice any vodka(*gen*) on the table"

It is important to note that, despite the genitive case, the noun *водки* "vodka" is clearly a direct object, since it is governed by a transitive verb. The genitive case

appears only due to negation, and in an affirmative sentence the use of genitive is not possible. Genitive is often used as *dobj* when it has semantics similar to the partitive case, and means 'some amount of the whole' [19]:

Rus: *Он дал мне денег* "He gave me some money(*gen*)"

Another use of genitive as direct object is referred to as intentional in [2]:

Rus: *Он ждал сигнала* "He waited for a signal(*gen*)"

In other words, the noun serves as an argument of intentional verbs, such as *to wait, to expect,* and *to search*, which normally take non-referential arguments. Such arguments do not correspond to a defined entity in the real world, and their existence is not clear (waiting for a signal does not necessarily mean that it will appear). Therefore, genitive case serves here to express this non-referentiality.

Finally, genitive forms appear also in prepositional object *pobj* when the preposition requires the genitive case:

Rus: *Я убежал от опасности* "I escaped from the danger(*gen*)"

Thus, the following rules were applied in order to handle different types of genitive objects:

1) *dobj* is used for genitive object with negation, intensional verbs, or similar to partitive

Rus: *Он искал поддержки* "He was looking for support(*gen*)"
dobj(искал, поддержки)

2) *pobj* is used for genitive object with preposition:

Rus: *Цитата из книги* "A quote from a book(*gen*)"
pobj(из, книги)

Another puzzling issue of Russian morpho-syntax is the genitive of quantification [21]. It appears that if the NP containing a numeral stands in a position which syntax assigns a direct case (nominative or accusative), the case of the noun is assigned by the numeral within the NP, regardless of the syntactic position:

Rus: *Я купил пять машин* "I bought five cars(*gen*)"
Rus: *Я купил машину* "I bought a car(*acc*)"

However, this does not happen when the NP is syntactically assigned and in the oblique case or in case of quantifiers:

Rus: *Я восхищаюсь пятью машинами* "I admire five cars(*instr*)"
Rus: *Я купил много машин* "I bought many car(*gen*)"

We treat these cases consistently, i.e. the noun is always the head of the numeral/quantifier, even if the case actually is assigned by the numeral/quantifier to the noun, e.g. *num*(машин, пять), *advmod*(машин, много).

2.6 Slavic Indirect Objects: *iobj* and Noun Head

The presence of cases in Slavic noun morphology creates some challenges for handling different types of objects within the SD model. The Russian language has 6 cases that are also common in other Slavic languages: nominative, accusative, genitive, dative, instrumental, and locative (Russian has almost lost the vocative case that still remains in other languages, e.g. Czech.)[4].

Nominative and accusative forms have corresponding dependency relations, which are covered in the initial version of the SD guidelines [7], i.e. nominal subject (*nsubj)* relation for nominative forms and direct object (*dobj)* for accusative forms. Locative case is used in Slavic languages mainly with prepositions, therefore it is analyzed using the *pobj* relation. Nouns in genitive case, when used as non-prepositional verb-objects, are mostly direct objects, or, if used as noun modifiers, they are mostly *gmod*. Thus, dative and instrumental objects are left uncovered, and it should be decided whether these objects need separate deprel tags as well, or if they can be merged under one generic indirect object (*iobj)* relation.

Since English only has dative indirect objects, the Russian *iobj* relation would be significantly different from the one defined in Stanford typed dependencies, where it is specific to the dative case: 'The indirect object of a VP[5] is the noun phrase which is the [dative] object of the verb' [7]:

She gave me a raise *iobj*(gave, me)

Moreover, there are a number of sentences in Russian where dative and instrumental indirect objects are both present, like in the following sentence:

Rus: *Она пишет ему письмо ручкой* "She writes him a letter with a pen"

Here, both *ему* "him" and *ручкой* "with a pen" will be indirect verb objects, even though they have different functions and correspond to different semantic roles, which can lead to syntactic ambiguity. Verbs that can take two indirect objects are, however, quite rare, and we decided to merge the two object relations into one in order to maintain cross-language consistency. Thus, dependency guidelines for Russian have the following definition for *iobj*: 'The indirect object of a VP is the

[4] Some researchers also distinguish so-called partitive genitive (*чашка чаю* "cup of tea"), and second locative cases (*в лесу* "in the forest").
[5] VP stands for a verb phrase. The VP has a verb as its head [29].

noun phrase without prepositions, which is the ablative, dative, or genitive object of the verb. This relation does not cover the genitive object with negated verbs and the genitive object denoting part of the whole that we treat as *dobj*).'

Another important issue related to Russian nominal cases arises from the fact that NPs in oblique cases can depend on other NPs. Consider the following phrase:

Rus: *письмо Лизе* "a letter to Liza(dat)"

As nouns do not normally take arguments, it is hard to say whether in these examples the nouns in oblique cases are objects or modifiers. This problem is sometimes referred to as 'argument-modifier ambiguity/distinction' [18]. Clear cases of noun argument occur in verb nominalization or in a noun with semantics of action, as in *удар кулаком* "a punch with a fist". Even though it does not fully comply with the initial definition of *iobj* and with the concept of object in general, we decided to treat these structures as indirect object relations in order to avoid creating a new deprel.

3 Slavic SD

In chapter 2, we proposed some modifications for the Slavic application of the SD schema. The main modifications are related to a new application of existing deprels, as in the case of *iobj* or *prt*. We also propose the addition of three new labels, two of them replacing some old labels. The addition of the deprel *gmod* is proposed for handling relatively frequent Slavic genitive modifier relations that do not have appropriate representation in the original SD schema. Original labels *attr*, *acomp,* and *cop* are replaced with labels *scomp* and *ocomp* in our model for improving consistency in labeling verbal complements. Also labels *abbrev* and *rel* are removed following proposals in [9]. Deprel *possessive* is removed, since possessive marker is not used in Slavic languages and this dependency relation is therefore obsolete. For deprel *ref*, we decided to analyze clause referents using deprels corresponding to the actual function of referent words in the sentence, so we analyze e.g. pronouns in relative clauses using the deprels *nsubj* or *dobj* depending on the internal head of the relative clause. Other deprels are used consistently with the original SD model [7], for the full set used in our model, see Table 1 below. The data annotated during the course of this project has become a part of the Universal Dependency Treebank (GSD[6]), and as such, some harmonization rules are applied [16]. The Universal SD (USD) presented in [10] brings some further cross-language consistency applications for multilingual treebanks. In order to compare between these related SD models and identify the differences between them, Table 1 also contains corresponding deprels used in GSD and USD. Our project-specific labels can be easily converted to match with the GSD or USD models, to achieve further consistency between different treebanks.

[6] We use GSD and USD as they are used in [10].

Table 1. List of deprels for Slavic SD application and corresponding GSD and USD deprels labels

Deprel	Gloss	GSD	USD
advcl	Adverbial clause modifier	advcl	advcl
advmod	Adverbial modifier	advmod	advmod
agent	Agent	adpmod	case
amod	Adjectival modifier	amod	amod
appos	Apposition	appos	appos
aux	Auxiliary verb	aux	aux
auxpass	Passive auxiliary	auxpass	auxpass
cc	Coordinating conjunction	cc	cc
ccomp	Clausal complement	ccomp	ccomp
complm	Clausal complement marker	mark	mark
conj	Conjunct	conj	conj
csubj	Clausal subject	csubj	csubj
csubjpass	Passive clausal subject	csubjpass	csubjpass
dep	Undetermined Dependent	dep	dep
det	Determiner	det	det
dobj	Direct object	dobj	dobj
emot	Emoticon	dep	dep
expl	Expletive	expl	expl
gmod	Genitive modifier	poss	poss
infmod	Infinitival modifier	infmod	nfincl
interj	Interjection	dep	dep
iobj	Indirect object	iobj	iobj
mark	Marker	mark	mark
mwe	Multi-word expression	mwe	mwe
neg	Negative particle	neg	neg
nn	Noun compound modifier	compmod	compound/name
npadvmod	NP adverbial modifier	nmod	nmod
nsubj	Nominal subject	nsubj	nsubj
nsubjpass	Passive nominal subject	nsubjpass	nsubjpass
num	Numeric modifier	num	nummod
number	Element of compound number	num	nummod
ocomp	Object complement	acomp/attr/cop	cop/xcomp
p	Punctuation	p	punct
parataxis	Parataxis	parataxis	parataxis
partmod	Participial modifier	partmod	nfincl
pcomp	Prepositional complement	adpcomp	ncmod
pobj	Prepositional object	adpobj	nmod
poss	Possession modifier	poss	poss
preconj	Preconjunct	cc	preconj
predet	Predeterminer	det	predet
prep	Prepositional modifier	adpmod	case
prt	Phrasal verb particle	prt	prt
purpcl	Purpose clause modifier	advcl	advcl
quantmod	Quantifier phrase modifier	advmod	advmod
rcmod	Relative clause modifier	rcmod	relcl
root	Root	root	root
scomp	Subject complement	acomp/attr/cop	cop/xcomp
tmod	Temporal modifier	advmod	advmod/tmod
xcomp	Open clausal complement	xcomp	xcomp

4 Conclusion

In this paper, we have presented an updated version of the SD schema for Slavic languages application. In order to keep cross-language consistency, we avoid introducing major changes to the schema that was already used for producing data resources for several languages. Proposed changes are applied using existing dependency labels for new language features that were not present in languages based on which the original SD schema was developed. Several new labels are introduced to improve consistency in handling verbal complements and in the case of *gmod*, the new label handles a language-specific feature that was not taken into account in the original SD model. In terms of consistency with existing SD treebanks, these new labels can be easily converted to labels used by other SD resources and vice versa. The updated SD schema expands coverage potential for SD treebanks and can be used both when building new SD based treebanks as well as for converting existing dependency treebanks into the SD schema.

References

1. Boguslavsky, I., Grigorieva, S., Grigoriev, N., Kreidlin, L., Frid, N.: Dependency treebank for Russian: Concept, tools, types of information. In: Proceedings of the 18th Conference on Computational Linguistics, vol. 2. Association for Computational Linguistics (2000)
2. Borschev, V., Partee, B., Paducheva, E., Testelets, Y., Yanovich, I.: Russian genitives, non-referentiality, and the property-type hypothesis (2007)
3. Bosco, C., Montemagni, S., Simi, M.: Converting Italian Treebanks: Towards an Italian Stanford Dependency Treebank. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (2013)
4. Buchholz, S., Marsi, E.: CoNLL-X shared task on multilingual dependency parsing. In: Proceedings of the 10th Conference on Computational Natural Language Learning. Association for Computational Linguistics (2006)
5. Chang, P., Tseng, H., Jurafsky, D., Manning, C.: Discriminative reordering with Chinese grammatical relations features. In: Proceedings of the Third Workshop on Syntax and Structure in Statistical Translation (SSST-3) at NAACL HLT (2009)
6. Comrie, B., Corbett, G.: The Slavonic Languages. Taylor & Francis (2002)
7. de Marneffe, M., Manning, C.: Stanford typed dependencies manual (2008)
8. de Marneffe, M., Connor, M., Silveira, N., Borschev, V.: Genitives, relational nouns, and the argument-modifier distinction. In: ZAS Papers in Linguistics, vol. 17, pp. 177–201 (2000)
9. de Marneffe, M., Connor, M., Silveira, N., Bowman, S.R., Dozat, T., Manning, C.: More constructions, more genres: Extending Stanford Dependencies. In: Proceedings of the Second International Conference on Dependency Linguistics, DepLing 2013 (2013)
10. de Marneffe, M., Dozat, T., Silveira, N., Haverinen, K., Ginter, F., Nivre, J., Manning, C.: Universal Stanford Dependencies: A Cross-linguistic typology. In: Proceedings of LREC (2014)
11. Downing, A., Locke, P.: A university Course in English Grammar. Prentice Hall (1992)
12. Džeroski, S., Erjavec, T., Ledinek, N., Pajas, P., Žabokrtský, Z., Žele, A.: Towards a Slovene dependency treebank. In: Proceedings of the 5th International Conference on Language Resources and Evaluation (2006)

13. Hajič, J., Vidová-Hladká, B., Pajas, P.: The Prague dependency treebank: Annotation structure and support. In: Proceedings of the IRCS Workshop on Linguistic Databases (2001)
14. Hajič, J., Ciaramita, M., Johansson, R., Kawahara, D., Martí, M., Màrquez, L., Meyers, A., Nivre, J., Padó, S., Štepánek, J., Straňák, P., Surdeanu, M., Xue, N., Zhang, Y.: The CoNLL-2009 shared task: syntactic and semantic dependencies in multiple languages. In: Proceedings of the 13th Conference on Computational Natural Language Learning. Association for Computational Linguistics, Boulder, Colorado (2009)
15. Haverinen, K., Viljanen, T., Laippala, V., Kohonen, S., Ginter, F., Salakoski, T.: Treebanking Finnish. In: Proceedings of TLT9, pp. 79–90 (2010)
16. McDonald, R., Nivre, J., Quirmbach-Brundage, Y., Goldberg, Y., Das, D., Ganchev, K., Hall, K., Petrov, S., Zhang, H., Täckström, O., Bedini, C., Bertomeu, C.N., Lee, J.: Universal Dependency Annotation for Multilingual Parsing. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (2013)
17. Nivre, J., Hall, J., Kübler, S., McDonald, R., Nilsson, J., Riedel, S., Yuret, D.: The CoNLL 2007 shared task on dependency parsing. In: Proceedings of EMNLPCoNLL (2007)
18. Partee, B., Borschev, V.: Genitives, relational nouns, and the argument-modifier distinction. In: ZAS Papers in Linguistics, vol. 17, pp. 177–201 (2000)
19. Paus, C.: Social and pragmatic conditioning in the demise of the Russian partitive case. Russian Linguistics 18(3), 249–266 (1994)
20. Petrov, S., Das, D., McDonald, R.: A universal part-of-speech tagset. In: Proceedings of LREC (2012)
21. Rappaport, G.: Numeral phrases in Russian: A minimalist approach. Journal of Slavic Linguistics 10(1-2), 327–340 (2002)
22. Seraji, M., Megyesi, B., Nivre, J.: Dependency Parsers for Persian. In: Proceedings of 10th Workshop on Asian Language Resources, COLING 2012, 24th International Conference on Computational Linguistics, Mumbai, India (2012)
23. Souček, M., Järvinen, T., LaMontagne, A.: Managing a Multilingual Treebank Project. In: Proceedings of the Second International Conference on Dependency Linguistics (2013)
24. Surdeanu, M., Johansson, R., Meyers, A., Màrquez, L., Nivre, J.: The CoNLL-2008 shared task on joint parsing of syntactic and semantic dependencies. In: Proceedings of CoNLL (2008)
25. Sussex, R., Cubberley, P.: The Slavic languages. Cambridge University Press, Cambridge (2006)
26. Tadić, M.: Building the Croatian Dependency Treebank: the initial stages. Suvremena Linguistica (2007)
27. Timberlake, A.: Reference conditions on Russian reflexivization. Language, 777–796 (1980)
28. Tsarfaty, R.: A Unified Morpho-Syntactic Scheme of Stanford Dependencies. In: Proceedings of ACL (2013)
29. Wardhaugh, R.: Understanding English Grammar: A Linguistic Approach. Blackwell (1995)
30. Wróblewska, A., Woliński, M.: Preliminary experiments in Polish dependency parsing. In: Bouvry, P., Kłopotek, M.A., Leprévost, F., Marciniak, M., Mykowiecka, A., Rybiński, H. (eds.) SIIS 2011. LNCS, vol. 7053, pp. 279–292. Springer, Heidelberg (2012)