

Synonymy in Bilingual Context: The CzEngClass Lexicon

Zdeňka Urešová Eva Fučíková Eva Hajičová Jan Hajič

Charles University

Faculty of Mathematics and Physics

Institute of Formal and Applied Linguistics

Malostranské nám. 25, CZ-11800 Prague, Czech Republic

{uresova, fucikova, hajicova, hajic}@ufal.mff.cuni.cz

Abstract

This paper describes CzEngClass, a bilingual lexical resource being built to investigate verbal synonymy in bilingual context and to relate semantic roles common to one synonym class to verb arguments (verb valency). In addition, the resource is linked to existing resources with the same or a similar aim: English and Czech WordNet, FrameNet, PropBank, VerbNet (SemLink), and valency lexicons for Czech and English (PDT-Vallex, Vallex and EngVallex). There are several goals of this work and resource: (a) to provide gold standard data for automatic experiments in the future (such as automatic discovery of synonym classes, word sense disambiguation, assignment of classes to occurrences of verbs in text, coreferential linking of verb and event arguments in text, etc.), (b) to build a core (bilingual) lexicon linked to existing resources, serving for comparative studies and possibly for training automatic tools, and (c) to enrich the annotation of a parallel treebank, the Prague Czech English Dependency Treebank, which so far contained valency annotation but has not linked synonymous senses of verbs together. The method used for extracting the synonym classes is a semi-automatic process with a substantial amount of manual work during filtering, role assignment to classes and individual class members' arguments, and linking to the external lexical resources. We present the first version with 200 classes (about 1800 verbs) and evaluate interannotator agreement using several metrics.

1 Introduction

Lexical resources, despite the fast progress in building end-to-end systems based on deep learning and artificial neural networks, are an important piece of the puzzle in Computational Linguistics and Natural Language Processing (NLP). They provide information that humans need to understand relations between words as well as the usage of these words in text. In addition, they can help various NLP tasks. This is why lexicons like WordNet (Miller, 1995; Fellbaum, 1998; Pala et al., 2011), FrameNet (Baker et al., 1998; Fillmore et al., 2003), VerbNet (Schuler, 2006), PropBank (Palmer et al., 2005) or EngVallex (Cinková, 2006; Cinková et al., 2014) have been created. They are for English, but there are also similar resources for other languages, often in a multilingual setting: WordNet is available in many languages (Vossen, 2004; Fellbaum and Vossen, 2012), Predicate Matrix (Lopez de Lacalle et al., 2016) extend coverage of several verbal resources and adds more romance languages, FrameNet has been extended to multiple languages (Boas, 2009), or there is the bilingual valency lexicon CzEngVallex for the case of Czech and English (Urešová et al., 2016).

One might thus question why it is necessary to develop another resource, and not just to connect some additional information to one or more of existing resources. After a thorough review (cf. also Sect. 2), it appears that for verbal synonymy, none of those resources fully corresponds to our goal of providing a lexicon of verbal synonyms based on at least semi-defined/semi-formal criteria, supported by real (parallel) texts usage. In addition, most of those resources have been first created for English and as some of the above publications acknowledge, there are then challenges to extend these resources to other languages (Fellbaum and Vossen, 2012).

This work is licensed under a Creative Commons Attribution 4.0 International License. License details: <http://creativecommons.org/licenses/by/4.0/>

In this paper, we describe CzEngClass, a lexical resource containing verbs in synonym groups (“classes”) that is based on certain criteria that help to determine membership in such classes. These criteria use information from valency lexicons and semantic frames, and rely on actual examples of use in a bilingual context. The lexicon is thus built “bottom up”, with emphasis on corpus evidence. In the resulting lexicon, links to existing lexical resources are added. Through these links, all entries also refer to examples from real bilingual texts with rich annotation, providing additional syntactic and morphological information for all CzEngClass entries.

The paper is structured as follows. In Sect. 2, related literature is discussed in more detail, while the resources actually used or referred to are described in Sect. 4. The criteria used for determining synonym class membership, specifically in bilingual context, are specified in Sect. 3. The resulting lexicon structure is presented in Sect. 5. Sect. 6 contains a description of the annotation process, while the interannotator agreement reached so far is analyzed in Sect. 7. We summarize some open questions and outline future work in Sect. 9.

2 Related work

Due to the needs of various NLP tasks, attention is paid to building interlinked resource(s) which integrates semantic information by mapping semantic knowledge from the individual lexical resources. One of the goals is to establish semantic interoperability between various semantic databases. As examples, we can name the SemLink (Palmer, 2009; Bonial et al., 2013) which provides manual mappings of four lexical resources (PropBank, VerbNet, FrameNet and WordNet), while the Predicate Matrix (Lopez de Lacalle et al., 2016) integrates and extends semantic information included in SemLink via automatic methods. BabelNet (Navigli and Ponzetto, 2012) integrates the largest multilingual Web encyclopedia(s), but mostly concentrates on “facts”. Besides a whole range of monolingual synonym lexicons such as Roget’s thesaurus (PSI and Associates, 1988) or WordNet (Miller, 1995; Fellbaum, 1998) for English, there are also cross-lingual synonym resources such as EuroWordNet, linking WordNet synsets across languages (Fellbaum and Vossen, 2012). WordNet contains a rich network of relations between its synsets (hyperonymy, hyponymy, ...), however, it does not contain information about syntactic and compositional semantic behavior, which is a drawback especially in case of verbs.

Our approach to verbal synonymy builds on previous research which also considered multilingual data (translations) in parallel corpora as an important resource. Translational context is regarded as a rich source of semantic information (de Jong and Appelo, 1987; Dyvik, 1998; Adamska-Sałaciak, 2013; Andrade et al., 2013). Parallel corpora have also been used (Resnik, 1997; Ide, 1999; Ide et al., 2002) for automatic methods for sense induction and disambiguation, for cross-lingual similarity detection and synonym extraction (Wu et al., 2010; Wu and Palmer, 2011). (Wang and Wu, 2012) studies various assumptions about synonyms in translation, for the purpose of trend detection from titles. While these works aim at automatic methods and applications, they share the idea that if two words are semantically similar in a language, their translations in another language would be also similar. Translations of a word from another language are often synonyms of one another (Lin et al., 2003; Wu and Zhou, 2003). A similar idea, i.e., that words sharing translational context are semantically related, can be found in (Plas and Tiedemann, 2006).

Interannotator agreement evaluation is regularly used in corpus annotation. However, it is much more scarcely used in lexicon entry creation and annotation, especially in a multilingual setting. For assignment of topics to words in Hindi and English, see (Kanojia et al., 2016). More detailed account on the influence of semantic lexical granularity within the Context Pattern Analysis paradigm on interannotator agreement can be found in (Cinková et al., 2012).

3 Synonymy in bilingual context

Synonymy in bilingual context is closely related to translational equivalence, sameness, similarity, meaning, word sense, etc. These terms - and the term synonymy itself - are not always used in an unambiguous way. We thus discuss the terminology first, and then specify how we define verbal synonymy in bilingual context in the work on building the CzEngClass lexicon.

3.1 Terminology

Although Cruse (1986) notes that “there is unfortunately no neat way of characterising synonyms”, the notion of synonymy is mostly seen as “sameness or identity of meaning” (Palmer, 1976; Sparck Jones, 1986). Leech (2012) restricts synonymy to equivalence of conceptual meaning. *Synonym* is mostly defined as “a same-language equivalent” (Adamska-Sałaciak, 2010; Adamska-Sałaciak, 2013) and “does not exceed the limits of a single language” (Gouws, 2013), while for bilingual contexts the term *translational equivalent* is used. On the other hand, (Martin, 1960; Klégr, 2004; Hahn et al., 2005; Hayashi, 2012; Haiyan, 2015; Dinu et al., 2015) recognize interlingual synonymy and use either the term *foreign-language equivalent*, *cross-lingual synonym*, *synonymous translation equivalent* or *bilingual synonym*.

In building CzEngClass, we consider the relationship between the lexical unit of the source language (SL) and of the target language (TL) unit as a specific type of synonymy, an *interlingual* synonymy. For words from different languages which are interlingual synonyms, we prefer to use the term *bilingual synonyms*. We understand that the meaning correspondence does not mean absolute equivalence and automatic interchangeability. We agree with (Louw, 2012) that the translation equivalent might be a TL item that can only be a substitute for the SL item in one or in some of its uses and each equivalent has to be supported with additional contextual and co-textual restrictions that will allow the user to make an appropriate choice of an “equivalent” for a given usage situation. Accordingly, we believe that interlingual synonymy considerations are essential in the translation process because it is up to the translator to choose the most suitable expression among (intralingual) synonyms, based on context and the meaning of the SL text (Catford, 1965; Newmark, 1988).

Intralingual synonyms are often not interchangeable, i.e., not quite equivalent. Synonymy, as viewed by (Lyons, 1968; Cruse, 1986), has to be seen as a scale of similarity (absolute, near and partial synonymy) and it is generally acknowledged that absolute synonymy is rare in natural languages. Therefore, it is believed that context must be taken into account (Palmer, 1981). Such synonyms are then called contextual synonyms (Zeng, 2007) or contextual correlates of synonymy (Rubenstein and Goodenough, 1965) and described as words that are “synonymously” used in certain specific texts.

3.2 Definition of contextual synonymy in CzEngClass

Both types of synonymy (interlingual and intralingual) are captured in the CzEngClass lexicon, which aims to group verbs into synonym classes both monolingually and cross-lingually. Along with (Palmer, 1981), we believe that synonymy can only be considered in context. We define two (or more) verb senses to be *bilingual synonyms* if they both (or all) convey the same meaning in a given particular context. Similarly, for the intralingual case, we work with the “loose” interpretation of synonymy (Lyons, 1968; Palmer, 1981), and consider context as a key factor that helps to overcome the vagueness of such “looseness”.

In CzEngClass, *context* is defined as the set of semantic roles (SRs) that the given verb, as a member of a bilingual synonym class, expresses by its arguments and/or adjuncts, or which are implicitly present, possibly with additional structural or semantic restrictions. Each class has an associated, single (common) set of SRs while such a set is shared by all its members, even if each SR can be expressed (*mapped* to) by a different argument (or by an adjunct, or implicitly or explicitly in the verb’s dependent substructure) for different verbs as members of that class. Conversely, such mapping must exist at least for all obligatory valency slots as defined in the two corresponding valency lexicons. To keep each class focused, only a relatively small set of SRs is usually assigned to it (corresponding roughly to “core” Frame Elements in FrameNet, even though the labels might not match).¹

Such focus then helps to answer the question of candidate verb membership in a particular synonym class. If a mapping of all arguments to the SRs associated with that class is found, as well as a mapping

¹In fact, there is one substantial difference between the intended properties of CzEngClass’s SRs and those found in FrameNet: while FrameNet explicitly states that SRs (FEs) from one Frame, even if under the same label, should not be construed as being the same as equally labeled SR (FE) in a different Frame, we would like to use such a set of SRs that *is* labeled consistently across CzEngClass classes. For the moment, however, we resort to VerbNet thematic roles for the most common SR “slots”, such as Agent and to a certain extent also Theme, renaming them in the future consistently once we have a larger set of classes ready. For more details on the process, see Sect. 6.

of all SRs of that class to the candidate verb arguments, adjuncts (or implicit, or even more deeply embedded dependents), and it does not violate any of the associated restrictions, then the candidate verb is considered to be a valid member of such class. Since CzEngClass is built mainly manually,² it is important to establish such criteria in order to achieve higher interannotator agreement (IAA) as well as for guiding the adjudication process. Table 1 illustrates one possible class and the context (SR mappings) for each member.

	Roles		
	Speaker_or_Event	Addressee	Content
povzbudit	ACT	ADDR	PAT
encourage	ACT	ADDR	PAT
galvanize	ACT	ADDR	PAT
inspire	ACT	ADDR	PAT
prod	ACT	ADDR	PAT
inspirovat	ACT	PAT	AIM
nabádat	ACT	ADDR	PAT
pobídnout	ACT	ADDR	PAT
podněcovat	ACT	ADDR	PAT
podpořit	ACT	ADDR	PAT
povzbuzovat	ACT	ADDR	PAT
vést	ACT	PAT	AIM

Table 1: Mappings for ENCOURAGE class

The following (parallel) example from the PCEDT illustrates the SR/argument use and alignment:

En: *Beth Marchand says Mrs. Yeargin.ACT/Speaker_or_Event inspired her.ADDR/Addressee to go.PAT/Content into education.*

Cz: *Beth Marchandová říká, že ji.PAT/Addressee učitelka Yearginová.ACT/Speaker_or_Event inspirovala, aby se dala.AIM/Content na studium vzdělávání.*

4 Resources used

The resources used to create the CzEngClass lexicon are divided into two groups - primary and secondary.

Primary resources are those used for word sense disambiguation and valency information when creating the class member candidates. We use the parallel Prague Czech-English Dependency Treebank 2.0 (PCEDT 2.0) (Hajič et al., 2012) with its associated lexicons. This treebank contains over 1.2 million words in almost 50,000 sentences for each language. About 90,000 tokens are verbs on each side. The English part contains the entire Penn Treebank - Wall Street Journal Section (Marcus et al., 1993). The Czech part is a manual translation of all the Penn Treebank-WSJ texts to Czech. PCEDT is annotated using The Prague Dependency Treebank style (Hajič et al., 2006; Hajič et al., 2018) manual linguistic annotation based on the Functional Generative Description framework (Sgall et al., 1986). Our research benefits primarily from the deep syntactico-semantic (tectogrammatical) dependency trees, interlinked across the two languages on sentence and node (content word) levels. Each deep syntax verb occurrence in the PCEDT is linked to the corresponding valency frame (predicate-argument structure frame) in the associated valency lexicons, PDT-Vallex (Urešová et al., 2014; Urešová, 2011) and EngVallex (Cinková, 2006; Cinková et al., 2014), effectively providing also word sense labeling. The parallel bilingual valency lexicon CzEngVallex (Urešová et al., 2016; Urešová et al., 2015), built over the PCEDT, is also heavily used, both in the annotation process itself but also for automatic preannotation.

Secondary resources are the well-known English lexical resources: FrameNet (Baker et al., 1998; Ruppenhofer et al., 2006), FrameNet+ (Pavlick et al., 2015), VerbNet and OntoNotes (Schuler, 2006; Pradhan et al., 2007), SemLink (Palmer, 2009; Bonial et al., 2012), PropBank (Palmer et al., 2005), and English WordNet (Miller, 1995; Fellbaum, 1998). For Czech, we use Czech WordNet (Pala and Smrž, 2004) and VALLEX (Lopatková et al., 2016). These resources are used for the extraction of an initial set of SRs (taken from FrameNet and VerbNet), and most importantly, their entries (if possible to the level of

²With automatic preselection of candidate class members based on parallel corpora and the valency lexicons available for Czech and English, see Sect. 6.

exact lexical units / frames / synsets) are referred to (explicitly linked) from all the corresponding entries in the CzEngClass lexicon.

5 Lexicon Structure

The structure of CzEngClass is described in detail (Urešová et al., 2018a; Urešová et al., 2018d; Urešová et al., 2017a; Urešová et al., 2017b). Here we summarize only its main characteristics.

The CzEngClass lexicon is in principle a set of (bilingual synonym) classes. Each class contains both Czech and English verbs (class members), identified by their valency frame identifier (i.e., a link to the Czech and English valency lexicons that served as the initial sense inventory for each verb). For each class, the lexicon records also the set of semantic roles, which is common for the class. For each class member, its arguments (valency slots) are mapped to this set (not necessarily 1 : 1 - arbitrary $m : n$ mappings are allowed). For each member, additional restrictions (for the time being, in the form of a plain text description) are recorded as well. In addition, the lexicon also records, within each class, the original verb pairs as found aligned across the two languages in the PCEDT, and their argument alignment as coming from the CzEngVallex bilingual valency lexicon. All external links (to FrameNet frame(s), Ontonotes Sense Groupings, VerbNet type(s), WordNet synset(s)) are also recorded (see Sect. 4).

The lexicon, technically a single XML file with external references, also contains a header with all the SRs used and their description, annotator IDs, bookkeeping information about all the external resources in order to create a URL for each external link, and all entries also contain the usual annotation information (log with timestamps, etc.).³

6 The lexicon annotation process

First, we automatically aligned Czech verbs with their English verbal translations⁴ as found in the PCEDT. There are two phases, each further broken down to several steps.

In Step 1 of the first phase, we have automatically extracted 200 Czech verbs⁵ and their English translations. These 200 verbs have been selected to represent both high-, medium- and low-frequency verbs in the PCEDT. They have been found in 23,769 sentences, covering about half of the corpus.⁶ Each Czech verb serves as a “seed” in the future bilingual synonym class. The English translations are the (first-phase, English-language) class member candidates. Using the principles of both intra-lingual and inter-lingual synonymy (Sect. 3), and with the help of comparing the valency frames of the Czech verb and the English verb and the English verbs among themselves, we then manually pruned these candidates, obtaining a list of (English) synonym class members.

In Step 2 of the first phase, we have used these English verbs and added, similarly to Step 1 but in an opposite direction and using the same corpus, additional Czech synonym class member candidates. These have been pruned manually again and a result of this pruning was a (first-phase) bilingual synonym class, built around the original Czech verb.

In the second phase, two tasks had to be carried out for all classes. First, to every class member the appropriate linking to English and Czech resources had to be provided. Using a dedicated class editor, we started with mapping every English verb sense in the class to Ontonotes Sense Grouping. Then, links to FrameNet, PropBank, WordNet and Czech VALLEX have been added.⁷ Second, core SRs inventory (for each class) had to be created. We were mostly inspired by FrameNet’s *frame elements* (FEs) resorting to more general VerbNet’s *thematic roles* if FEs were deemed to be specific. For each class member, its valency arguments (taken from PDT-Vallex for Czech verbs and EngVallex for English ones) have been mapped to the appropriate SR from the set of SRs assigned to the whole class. So far, this second step has been performed for the first 60 classes only.

³Current version of CzEngClass: <http://hdl.handle.net/11234/1-2824>

⁴As already mentioned, we pay attention only to verbs.

⁵I.e., 200 verb senses represented by their valency frames as annotated in the data.

⁶Average number of sentences per Czech verb is 119, median is 26. These sentences may contain also other verbs.

⁷Links to Czech WordNet will be added later.

The process described above was first tested and fine-tuned by two annotators, and then a detailed testing (so far, of Step 1 of the first phase), as reported in Sect. 7 below, has been performed by eight annotators.

7 Evaluation: interannotator agreement

Using the first part of the annotation process (filtering of the precomputed list of English verbs for a given Czech verb, in a particular sense), an IAA evaluation scheme has been set up (Table 2).

	No. of classes (Czech verbs)	No. of English verbs (sum over all classes)
Part 1	15	131
Part 2	15	134
Part 3	17	127
Part 4	13	152
Total	60	544

Table 2: Interannotator agreement experiment setup

As described in Sect. 6, the 60 Czech verbs (senses) have been selected in the PCEDT such that they represented high-, medium- and low-frequency verbs in roughly the same proportion.⁸ For those 60 verbs, all their English translational verbal counterparts⁹ aligned with them have been extracted automatically from the corpus, with the help of the CzEngVallex lexicon which refers to particular verb senses. This alignment is however not perfect for our purposes. The task of the annotators was to prune, for each Czech verb sense (i.e., the potential/future synonym class), the automatically preselected list of its English verbal translational equivalents so that only synonym candidates corresponding to the meaning of the Czech verbs as it appears in the bilingual corpus are preserved. For this experiment, the annotators have not been given the full definition of the synonym as presented here in Sect. 3. They have been instructed, for each pair of the Czech verb and its English synonym candidate, to check the corpus examples through the CzEngVallex valency argument alignments, i.e., the usage in context as defined in Sect. 3, but not to validate (yet) against the set of SRs for the given class.¹⁰

There have been 8 annotators, each of them annotating 2 different parts. 1 of the annotators is a professional translator, 2 of them are researchers in the field of Computational linguistics with a Ph.D. degree, and the remaining 5 are undergraduate students of various language and linguistic programs, all with high-level fluency in English. All annotations have been treated as equal.

Each English synonym class member candidate verb was annotated by 4 annotators (Table 2). The annotators have been asked to assign, to each candidate English verb, one of the five labels summarized in Table 3, corresponding to a scaled decision of how strong the annotators feel about including (or not including) the English word in the resulting synonym group.

Only the labels have been presented to the annotators, which have been then converted to the scale used for some of the evaluation metrics, as described below. The data for the interannotator agreement as described in this paper are available with the previous version of CzEngClass (Urešová et al., 2018b).¹¹

7.1 Interannotator agreement: Cohen’s kappa

Cohen’s kappa, used for comparing annotations pairwise, is defined as follows:

$$\kappa = (p_0 - p_e) / (1 - p_e) \quad (1)$$

⁸This selection is to gain insight to the issues with verbs of all frequencies; overall, of course, the verbs with low frequency will prevail, making the statistics more favorable since classes for these verbs tend to have less members and thus higher IAA.

⁹We are aware of the fact that many verbs are translated by nouns or other kinds of phrases. These translational counterparts are not yet part of the CzEngClass.

¹⁰We understand this as a first step. The class membership might be—and surely will be—changed in the second step, namely, when assigning the SR mappings to valency slots. This is the most important “context” criterion defining the class membership based on our principles of synonymy as defined in Sect. 3, and specifically in Sect. 3.2.

¹¹<http://hdl.handle.net/11234/1-2823>

Label	Explanation	Inclusion weight
Y	Yes	4
R _Y	Rather Yes	3
R _N	Rather No	2
N	No	1
D	Delete (Alignment or other system error)	0

Table 3: Labels assigned to each English verb by the annotators

where p_0 is the observed agreement, and p_e the expected agreement based on the two annotations. Results are summarized in Table 4. For determining agreement, we have mapped the labels assigned by the annotators to two values only: Y (if they used labels Y and R_Y) and N (if they used labels N, R_N and D), since Cohen’s kappa would be hard to interpret if scaled values are behind the labels used.

Annotator	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈
A ₁	-	0.46	0.19	0.42		0.43	0.43	0.33
A ₂		-	0.57	0.55	0.46	0.68		0.31
A ₃			-	0.42	0.37		0.26	
A ₄				-	0.35		0.45	
A ₅					-	0.48	0.56	0.41
A ₆						-	0.48	0.51
A ₇							-	0.48
A ₈								-

Table 4: Interannotator agreement: Cohen’s kappa, pairwise

Due to the distribution of Parts 1-4 to annotators, not all pairs of annotators could be compared (empty cells in Table 4). The parts in any row or column are different for different annotators.¹²

While averaging over the annotators is not well defined within such pairwise distribution of data, we can see quite clearly that A₃ has overall the lowest agreement with the others, but the differences are not that high. The values are low, but this is due to the fact that the label assignment has been biased to assigning Y (i.e., to keep the verb in) in about an 5:1 ratio.

7.2 Interannotator agreement: Fleiss’ kappa

As opposed to Cohen’s kappa, Fleiss’ kappa can be used in a multiannotator setup with n annotators and N datapoints (total of English verbs in the 60 classes annotated, i.e., 544 in our case):

$$\kappa = (\bar{P} - \bar{P}_e) / (1 - \bar{P}_e) \quad (2)$$

where

$$\bar{P} = \frac{1}{N} \sum_{i=1}^N \frac{1}{n(n-1)} (n_{iY} (n_{iY} - 1) + n_{iN} (n_{iN} - 1)), \quad (3)$$

$\bar{P}_e = p_Y^2 + p_N^2$, $p_Y = \frac{1}{Nn} \sum_{i=1}^N n_{iY}$ and $p_N = \frac{1}{Nn} \sum_{i=1}^N n_{iN}$. The n_{iY} (n_{iN}) counts are defined as the number of times the i -th datapoint has been annotated Y and N, respectively (i.e., sum of n_{iY} and n_{iN} is n). The Ys and Ns have been obtained by the same mapping as in the Cohen’s kappa case: Y and R_Y mapped to Y and N, R_N and D mapped to N. Table 5 shows the intermediate and final values of Fleiss’ kappa. The 0.45 value characterizes the “global” difficulty of the task; again, we have to stress that the bias is high (about 5:1), so that any disagreement lowers the kappa value substantially.

¹²Given the amount of effort available for the annotation, the four parts have been distributed in such a way that at least those pairs listed in the table could be covered.

p_Y	p_N	\bar{P}	\bar{P}_e	κ
0.84	0.16	0.85	0.73	0.45

Table 5: Interannotator agreement: Fleiss’ kappa, global

Given the low kappa values, we have been interested in numbers that could perhaps be easier to interpret from the point of view of adjudication, which will inevitably follow to ensure high quality of the resulting lexicon. In addition, we are not convinced that these widely used interannotator metrics are appropriate for highly biased lexical task (as opposed for running text annotation in a corpus). We have thus measured two more things: consensus and deviation from the assumed correct value, both described below.

7.3 Annotator consensus

As a first step, we have measured consensus. Of the 544 datapoints the annotators assigned the label (again mapped to Y and N only), in 488 cases a majority consensus could be reached (89.7%); i.e., only in slightly over 10% of the cases the result was 2:2 (with 4 annotators working on each datapoint). This number naturally varies with the number of annotators (it would be higher for 2 annotators, and always 100% for an odd number of annotators unless a higher threshold is set for determining consensus), but when combined with full agreement among the 4 annotators (358 cases out of the 544, or 65.8%), it gives a good idea of the adjudication effort needed (for 4 annotators, in our case).¹³

7.4 Comparison based on the annotated scale

Since the annotation of class membership for the candidate English synonym verbs has been in fact on a scale 0-4 (see Table 3), we tried to compute a global (averaged) score of deviation from a centerpoint, namely the assumed correct value, computed either as a majority or average value as assigned to every English verb by the four different annotators that had labelled it. We have used two methods.

In the **first method**, we have determined the “correct” value by taking the majority annotation, restricted again to two values: Y(es) and N(o) (with mapping from the five different (scaled) annotated labels as in the kappa computation, Sect. 7.1). For computing the deviation on a particular datapoint, we use the same mapping. In other words, the deviation can only be 0 (agreement with Y or N) or 3 (disagreement). The averaged (absolute) deviation $DevA2_j$ for an annotator j is then defined as

$$DevA2_j = \frac{\sum_{i=1}^N |B_i - S_{ij}|}{N} \quad (4)$$

where B_i is the correct value (4 for Y, 1 for N) for verb i (out of N verbs, where $N = 544$ in our dataset), and S_{ij} is the mapped value on that i -th verb for annotator j .

The deviation computed on this mapped (Y/N) values is summarized for all annotators in Table 6; the average over all annotators is 0.27 (within a span of maximum possible difference of 3 (4-1)). Since the bias towards the Y label in the consensus labeling is about 5:1, giving everything Y would result in approx. average deviation $1/6 \times 3 = 0.5$, so the annotation cut this deviation to half.

Annotator	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈	Data avg.
Avg. deviation	0.21	0.20	0.23	0.26	0.41	0.21	0.31	0.31	0.27

Table 6: Deviation from consensus value, mapping to 2 values only (Y/N)

The **second method**, which is meant for evaluation purposes only, does not select one consensus value, but simply averages the scaled labels over all (four) labels assigned by the annotators to every datapoint (verb to annotate) i :

¹³The total agreement has been measured using all labels; i.e., if 3 annotators assigned N and 1 assigned R_N, it did not count as full agreement.

$$P_i = (\sum_{k=1}^n v_{ik})/n$$

where n is a number of annotations for a given datapoint (i.e., always four in our case, since each verb has been annotated by 4 annotators), k runs over all labels assigned to the i -th datapoint, and v_{ik} is the k -th value (0-4, mapped from (D, N, R_N, R_Y, Y; see Table 3) assigned to i -th datapoint.

The deviation for each annotator is then computed as

$$DevA5_j = \frac{\sum_{i=1}^N |P_i - A_{ij}|}{N} \quad (5)$$

where P_i is the averaged value as defined above for verb i (out of N verbs, where $N = 544$ in our dataset), and A_{ij} is value assigned to the i -th datapoint by annotator j .

The averaged values are shown in Table 7.

Annotator	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈	Data avg.
Avg. deviation	0.32	0.32	0.32	0.32	0.45	0.36	0.37	0.41	0.36

Table 7: Deviation from average assigned value, all 5 values used (0-4)

These deviations are (somewhat surprisingly) larger than the deviations computed after mapping the labels to the binary (Y/N) values only. In other words, the deviation is *smaller* when the “unsure” labels (R_Y, R_N) are mapped to Y and N, and the D value also to N, disregarding the difference between these two.

Since the final lexicon will not contain any manually assigned weights, i.e., a given verb will be considered either in a class or not in a class. we believe that the relevant numbers are those in Table 6.

While the two methods of computing interannotator “deviation” compute absolute difference between the annotator-assigned value and the “truth”, we have been interested yet in another value, namely whether a given annotator tends to keep the preselected verbs in the class or not. This number thus does not describe annotator’s distance from the consensus value or the average, but rather his average positive attitude (prefers to keep many verbs in) or a negative one (prefers to mark verbs as not belonging to the synonym class in question).

We compute the “truth” as a simple average of the graded labels for every verb:

$$DevS5_j = \frac{\sum_{i=1}^N (P_i - A_{ij})}{N} \quad (6)$$

The resulting value can now be also negative, meaning that the annotator assigned lower values on the 0-4 scale than average, and by how much.

Annotator	A ₁	A ₂	A ₃	A ₄	A ₅	A ₆	A ₇	A ₈
Avg. (signed) deviation	0.09	0.05	0.13	0.08	-0.19	0.07	-0.02	0.11

Table 8: Signed deviation from average assigned value, all 5 values used (0-4)

Table 8 shows the results. The deviation for 2 annotators, A₅ and A₇ is in the negative territory. These annotators, and specifically A₅, are considered “conservative”, since their annotation keeps the synonym classes small - they prefer only clear synonyms to become included in the class. Annotator A₃, on the other hand, is the most “liberal” of all and leans towards including more words in the class.

8 Extending CzEngClass to Other Languages

An interesting question is what is necessary in order to create a similar lexical resource—using the approach described here—for other languages or to extend our resource by another language. Based on our

experience so far, we believe this is the minimal set of necessities to start with: a valency or predicate-argument lexicon for that language (a corresponding parallel resource - with English, such as our CzEng-Vallex lexicon - might be helpful but it is not essential) and a word-aligned and (at least automatically) syntactically analyzed parallel corpus (pairing the language in question to English, or to any language that is already included in the Class lexicon being extended). In addition, some word-sense-distinguishing lexicon (e.g., WordNet in that language) is also nice to have to link the class members to. Moreover, having additional lexical resources in the given language (such as FrameNet) would be an advantage that will help the consistency of annotation and the usefulness of the final, externally linked synonym lexicon, but this is not strictly necessary.

9 Conclusions and future work

We have described the first version of a semi-manually built bilingual synonym lexicon of verbs, CzEng-Class, based on a parallel corpus and existing lexical resources. The process of building the lexicon is relatively complex, and we have presented results of the first step of “soft” (scaled) lexical unit annotation by several annotators. We have evaluated their agreement by several metrics, giving us some indication of how much future effort will be needed for adjudication in order to obtain a high-quality resource.

These first findings indicate that the traditional measures for IAA show low agreement, primarily due to high bias in the automatic preselection which favors the preselected verbs to be kept (in 5:1 ratio).¹⁴ Using a different measure based on deviation from consensus, we found that a full consensus has been achieved among 4 annotators in almost 2/3rd of test cases, and if we lower the criterion of agreement to 3 of 4 annotators, in 9 of 10 cases (Sect. 7.3). This is only when considering the verb sense and valency; adding the semantic roles, which is the next step, would likely lower that agreement. For example, we will explore more the assumption (and try to explicitly define where and when it can be used) that the valency slots to semantic roles mapping might need to be $m:n$, which of course greatly influences the interannotator agreement as opposed to the case where only $1:1$ mapping is allowed.

The first version (200 classes, i.e., approx. 1800 class members, Phase 1 annotation) has been released publicly (Urešová et al., 2018c).¹⁵

Phase 2 annotation is ongoing, with semantic role assignment to classes and the detailed mapping of valency to these roles. Eventually, we will extend CzEngClass to all verbs found in the PCEDT. This is expected to go smoothly, perhaps with the usual exception of the verbs *to be*, *to have*, *to become* and *to do*. Otherwise, since we have already (intentionally) dealt with verbs of various frequencies and degrees of homonymy, representing a wide range of possible valency-to-semantic-roles mappings, we believe no substantial issues in the specification nor the annotation process should arise.

For the more distant future, we plan to use the core PCEDT-derived CzEngClass lexicon to do experiments on automatic extraction of synonym classes from other (larger) data, while using CzEngClass as a gold standard. If successful, we plan to extend CzEngClass to other verbs similarly to what the Predicate Matrix project did, but with the arguably best system based on the above experiments; in all cases, we plan for manual pass to obtain consistent classes with those done fully manually.

Acknowledgements

This work has been supported by the grant No. GA17-07313S of the Grant Agency of the Czech Republic. Part of this work was conducted during a fellowship (of the 4th author) at the Oslo Center for Advanced Study at the Norwegian Academy of Science and Letters. The data used in this work have been created and are maintained in the LINDAT/CLARIN digital repository (<http://lindat.cz>), supported by the Ministry of Education, Youth and Sports of the Czech Republic as projects No. LM2015071 and CZ.02.1.01/0.0/0.0/16_013/0001781. The creation of part of the data used has been also supported by the European Commission under the T4ME (META-NET) project No. FP7-ICT-2009-4-249119.

¹⁴Which confirms, as a collateral result, that the data from which this preselection has been done are of quite high quality in terms of verb and argument alignment.

¹⁵<http://hdl.handle.net/11234/1-2824> available under CC BY-NC-SA 4.0.

References

- Arleta Adamska-Sałaciak. 2010. Examining equivalence. *International Journal of Lexicography*, 23(4):387–409.
- Arleta Adamska-Sałaciak. 2013. Equivalence, Synonymy, and Sameness of Meaning in a Bilingual Dictionary. *International Journal of Lexicography*, 26:329.
- Daniel Andrade, Masaaki Tsuchida, Takashi Onishi, and Kai Ishikawa. 2013. Synonym acquisition using bilingual comparable corpora. In *Sixth International Joint Conference on Natural Language Processing, IJCNLP 2013, Nagoya, Japan, October 14-18, 2013*, pages 1077–1081.
- Collin F. Baker, Charles J. Fillmore, and John B. Lowe. 1998. The Berkeley FrameNet Project. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics - Volume 1, ACL '98*, pages 86–90, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Hans C. Boas, editor. 2009. *Multilingual FrameNets in Computational Lexicography. Methods and Applications*. Trends in Linguistics. Studies and Monographs [TiLSM] 200. DE GRUYTER MOUTON.
- Claire Bonial, Weston Feely, Jena D Hwang, and Martha Palmer. 2012. Empirically Validating VerbNet Using SemLink. In *Seventh Joint ACL-ISO Workshop on Interoperable Semantic Annotation*, Istanbul, Turkey, May.
- Claire Bonial, Kevin Stowe, and Martha Palmer. 2013. *Proceedings of the 2nd Workshop on Linked Data in Linguistics (LDL-2013): Representing and linking lexicons, terminologies and other language data*, chapter Renewing and Revising SemLink, pages 9 – 17. Association for Computational Linguistics.
- John Cunnison Catford. 1965. *A Linguistic Theory of Translation: an Essay on Applied Linguistics*. Oxford Univ. Press, London.
- Silvie Cinková, Martin Holub, and Vincent Kríž. 2012. Optimizing semantic granularity for NLP - report on a lexicographic experiment. In Ruth Fjeld and Julie Torjusen, editors, *Proceedings of the 15th EURALEX International Congress*, number hardcopy ISBN: 978-82-303-2095-2, pages 523–531, Oslo, Norway. Department of Linguistics and Scandinavian Studies, University of Oslo, Department of Linguistics and Scandinavian Studies, University of Oslo.
- Silvie Cinková, Eva Fučíková, Jana Šindlerová, and Jan Hajič. 2014. *EngVallex - English Valency Lexicon*. LINDAT/CLARIN digital library. <http://hdl.handle.net/11858/00-097C-0000-0023-4337-2>.
- Silvie Cinková. 2006. From PropBank to EngVallex: Adapting the PropBank-Lexicon to the Valency Theory of the Functional Generative Description. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006)*, pages 2170–2175, Genova, Italy. ELRA.
- D.Alan Cruse. 1986. *Lexical Semantics*. Cambridge University Press, UK.
- Franciska M.G. de Jong and Lisette Appelo. 1987. Synonymy and translation, 12.
- Anca Dinu, Liviu P. Dinu, and Ana Sabina Uban. 2015. Cross-lingual synonymy overlap. In *Recent Advances in Natural Language Processing, RANLP 2015, 7-9 September, 2015, Hissar, Bulgaria*, pages 147–152.
- Helge Dyvik. 1998. Translations as Semantic Mirrors: From Parallel Corpus to Wordnet. In *Proceedings of Workshop Multilinguality in the Lexicon II, ECAI 98*.
- Christiane Fellbaum and Piek Vossen. 2012. Challenges for a multilingual wordnet. *Lang. Resour. Eval.*, 46(2):313–326, June.
- Christiane Fellbaum. 1998. *WordNet: An Electronic Lexical Database*. Language, Speech, and Communication. MIT Press, Cambridge, MA. 423 pp.
- Charles J. Fillmore, Ch. R. Johnson, and M. R. L.Petruck. 2003. Background to FrameNet: FrameNet and Frame Semantics. *International Journal of Lexicography*, 16(3):235–250.
- Rufus H. Gouws. 2013. Contextual and co-textual guidance regarding synonyms in general bilingual dictionaries. *International Journal of Lexicography*, 26(3):346–361.
- Udo Hahn, Philipp Daumke, Stefan Schulz, and Kornél G. Markó. 2005. Cross-language mining for acronyms and their completions from the web. In *Discovery Science, 8th International Conference, DS 2005, Singapore, October 8-11, 2005, Proceedings*, pages 113–123.

- Men Haiyan. 2015. *Vocabulary increase and collocation learning: a corpus-based cross-sectional study of Chinese EFL learners*. Ph.D. thesis, Birmingham City University.
- Jan Hajič, Eva Hajičová, Jarmila Panevová, Petr Sgall, Silvie Cinková, Eva Fučíková, Marie Mikulová, Petr Pajas, Jan Popelka, Jiří Semecký, Jana Šindlerová, Jan Štěpánek, Josef Toman, Zdeňka Urešová, and Zdeněk Žabokrtský. 2012. *Prague Czech-English Dependency Treebank 2.0*. <https://catalog.ldc.upenn.edu/LDC2004T25>.
- Jan Hajič, Eduard Bejček, Alevtina Bémová, Eva Buráňová, Eva Hajičová, Jiří Havelka, Petr Homola, Jiří Kárník, Václava Kettnerová, Natalia Klyueva, Veronika Kolářová, Lucie Kučová, Markéta Lopatková, Marie Mikulová, Jiří Mírovský, Anna Nedoluzhko, Petr Pajas, Jarmila Panevová, Lucie Poláková, Magdaléna Rysová, Petr Sgall, Johanka Spoustová, Pavel Straňák, Pavlína Synková, Magda Ševčíková, Jan Štěpánek, Zdeňka Urešová, Barbora Vidová Hladká, Daniel Zeman, Šárka Zikánová, and Zdeněk Žabokrtský. 2018. *Prague Dependency Treebank 3.5*. LINDAT/CLARIN digital library. <http://hdl.handle.net/11234/1-2621>.
- Jan Hajič, Jarmila Panevová, Eva Hajičová, Petr Sgall, Petr Pajas, Jan Štěpánek, Jiří Havelka, Marie Mikulová, Zdeněk Žabokrtský, Magda Ševčíková Razímová, and Zdeňka Urešová. 2006. *Prague Dependency Treebank 2.0*. Number LDC2006T01. LDC, Philadelphia, PA, USA. <https://catalog.ldc.upenn.edu/ldc2006t01>.
- Yoshihiko Hayashi. 2012. Computing Cross-lingual Synonym Set Similarity by Using Princeton Annotated Gloss Corpus. In *GWC 2012: 6th International Global Wordnet Conference, Proceedings*, pages 134–141. Tribun EU s. r. o.
- Nancy Ide, Tomaz Erjavec, and Dan Tufis. 2002. Sense discrimination with parallel corpora. In *Proceedings of the ACL-02 Workshop on Word Sense Disambiguation: Recent Successes and Future Directions*.
- Nancy Ide. 1999. Parallel translations as sense discriminators. In *SIGLEX99: Standardizing Lexical Resources*.
- Diptesh Kanojia, Aditya Joshi, Pushpak Bhattacharyya, and Mark James Carman. 2016. That'll do fine!: A coarse lexical resource for english-hindi mt, using polylingual topic models. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, Paris, France, may. European Language Resources Association (ELRA).
- Aleš Klégr. 2004. Vnitrojazyková a mezijazyková synonymie (Intralingual and Interlingual Synonymy), philologica.net, ISSN 1214-5505, available online. (2004-01-13).
- Geoffrey N. Leech, Marianne Hundt, Christian Mair, and Nicholas Smith. 2012. *Change in Contemporary English*. Cambridge University Press, New York.
- Dekang Lin, Shaojun Zhao, Lijuan Qin, and Ming Zhou. 2003. Identifying synonyms among distributionally similar words. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence, IJCAI'03*, pages 1492–1493, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.
- Markéta Lopatková, Václava Kettnerová, Eduard Bejček, Anna Vernerová, and Zdeněk Žabokrtský. 2016. *Valenční slovník českých sloves VALLEX*. Nakladatelství Karolinum, Praha, Czechia.
- Maddalen Lopez de Lacalle, Egoitz Laparra, Itziar Aldabe, and German Rigau. 2016. Predicate matrix: automatically extending the semantic interoperability between predicate resources. *Language Resources and Evaluation*, 50(2):263–289, Jun.
- Phillip Adriaan Louw. 2012. Synonymy in the translation equivalent paradigms of a standard translation dictionary. *Lexikos*, 8(1).
- John Lyons. 1968. *Introduction to Theoretical Linguistics*. Cambridge University Press.
- Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. 1993. Building a Large Annotated Corpus of English: The Penn Treebank. *COMPUTATIONAL LINGUISTICS*, 19(2):313–330.
- Samuel E. Martin. 1960. *Selection and presentation of ready equivalents in a translation dictionary, work paper for conference on lexicography, Indiana university, November 11-12, 1960 [microform] / Samuel E. Martin*. Distributed by ERIC Clearinghouse [Washington, D.C.].
- George A. Miller. 1995. WordNet: A Lexical Database for English. *Commun. ACM*, 38(11):39–41, November.

- Roberto Navigli and Simone Paolo Ponzetto. 2012. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence*, 193:217–250.
- Peter Newmark. 1988. *A textbook of translation*. Prentice-Hall International, New York.
- Karel Pala and Pavel Smrž. 2004. Building Czech Wordnet. *Romanian Journal of Information Science and Technology*, 7(1-2):79–88.
- Karel Pala, Tomáš Čapek, Barbora Zajíčková, Dita Bartůšková, Kateřina Kulková, Petra Hoffmannová, Eduard Bejček, Pavel Straňák, and Jan Hajič. 2011. *Czech WordNet 1.9 PDT*. LINDAT/CLARIN digital library. <http://hdl.handle.net/11858/00-097C-0000-0001-4880-3>.
- Martha Palmer, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71–106, March.
- Frank Robert Palmer. 1976. *Semantics: A New Outline*. Cambridge University Press.
- Frank Robert Palmer. 1981. *Semantics. 2nd edn*. Cambridge University Press, UK.
- Martha Palmer. 2009. SemLink: Linking PropBank, VerbNet and FrameNet. In *Proceedings of the Generative Lexicon Conference*, page 9–15.
- Ellie Pavlick, Travis Wolfe, Pushpendre Rastogi, Chris Callison-Burch, Mark Dredze, and Benjamin Van Durme. 2015. Framenet+: Fast paraphrastic tripling of framenet. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 408–413, Beijing, China, July. Association for Computational Linguistics.
- Lonneke Plas and Jörg Tiedemann. 2006. Finding synonyms using automatic word alignment and measures of distributional similarity. In *Proceedings of the COLING/ACL 2006 Main Conference Poster Sessions*, pages 866–873. Association for Computational Linguistics.
- Sameer S. Pradhan, Eduard Hovy, Mitch Marcus, Martha Palmer, Lance Ramshaw, and Ralph Weischedel. 2007. Ontonotes: A unified relational semantic representation. *International Journal of Semantic Computing*, 01(04):405–419.
- PSI and Associates. 1988. *Roget's thesaurus of synonyms and antonyms*. P.S.I. & Associates.
- Philip Resnik. 1997. A perspective on word sense disambiguation methods and their evaluation. In *Tagging Text with Lexical Semantics: Why, What, and How?*
- Herbert Rubenstein and John B. Goodenough. 1965. Contextual correlates of synonymy. *Commun. ACM*, 8(10):627–633, October.
- J. Ruppenhofer, M. Ellsworth, M. R. L. Petruck, Ch. R. Johnson, and J. Scheffczyk. 2006. FrameNet II: Extended theory and practice. *Unpublished Manuscript*.
- Karin Kipper Schuler. 2006. *VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon*. Ph.D. thesis, University of Pennsylvania.
- Petr Sgall, Eva Hajičová, and Jarmila Panevová. 1986. *The meaning of the sentence in its semantic and pragmatic aspects*. D. Reidel, Dordrecht.
- Karen Sparck Jones. 1986. *Synonymy and Semantic Classification*. Edinburgh University Press, Edinburgh, Scotland, Scotland.
- Zdeňka Urešová, Jan Štěpánek, Jan Hajič, Jarmila Panevová, and Marie Mikulová. 2014. *PDT-Vallex*. LINDAT/CLARIN digital library. <http://hdl.handle.net/11858/00-097C-0000-0023-4338-F>.
- Zdeňka Urešová, Eva Fučíková, Jan Hajič, and Jana Šindlerová. 2015. *CzEngVallex - Czech-English Valency Lexicon*. LINDAT/CLARIN digital library. <http://hdl.handle.net/11234/1-1512>.
- Zdeňka Urešová, Eva Fučíková, and Jana Šindlerová. 2016. CzEngVallex: a bilingual Czech-English valency lexicon. *The Prague Bulletin of Mathematical Linguistics*, 105:17–50.
- Zdeňka Urešová, Eva Fučíková, and Eva Hajičová. 2017a. Czengclass – towards a lexicon of verb synonyms with valency linked to semantic roles.

- Zdeňka Urešová, Eva Fučíková, Eva Hajičová, and Jan Hajič. 2017b. Syntactic-semantic classes of context-sensitive synonyms based on a bilingual corpus. In Zygmunt Vetulani and Joseph Mariani, editors, *Proceedings of 8th Language and Technology Conference*, pages 201–205, Poznań, Poland. Fundacja Uniwersytetu im. Adama Mickiewicza, Fundacja Uniwersytetu im. Adama Mickiewicza w Poznaniu.
- Zdeňka Urešová, Eva Fučíková, Eva Hajičová, and Jan Hajič. 2018a. Creating a Verb Synonym Lexicon Based on a Parallel Corpus. In *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC'18)*, Miyazaki, Japan, May. European Language Resources Association (ELRA).
- Zdeňka Urešová, Eva Fučíková, Eva Hajičová, and Jan Hajič. 2018b. CzEngClass 0.1. LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, <http://hdl.handle.net/11234/1-2823>, Creative Commons - Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0).
- Zdeňka Urešová, Eva Fučíková, Eva Hajičová, and Jan Hajič. 2018c. CzEngClass 0.2. LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, <http://hdl.handle.net/11234/1-2824>, Creative Commons - Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0).
- Zdeňka Urešová, Eva Fučíková, Eva Hajičová, and Jan Hajič. 2018d. Tools for Building an Interlinked Synonym Lexicon Network. In *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC'18)*, Miyazaki, Japan, May. European Language Resources Association (ELRA).
- Zdeňka Urešová. 2011. *Valenční slovník Pražského závislostního korpusu (PDT-Vallex)*. Studies in Computational and Theoretical Linguistics. Ústav formální a aplikované lingvistiky, Praha, Czechia. 375 pp.
- Piek Vossen. 2004. Eurowordnet: A multilingual database of autonomous and language-specific wordnets connected via an inter-lingualindex. *International Journal of Lexicography*, 17(2):161–173.
- Fei Wang and Yunfang Wu. 2012. Mining market trend from blog titles based on lexical semantic similarity. In *Computational Linguistics and Intelligent Text Processing - 13th International Conference, CICLing 2012, New Delhi, India, March 11-17, 2012, Proceedings, Part II*, pages 261–273.
- Shumin Wu and Martha Palmer. 2011. Semantic mapping using automatic word alignment and semantic role labeling. In *Proceedings of Fifth Workshop on Syntax, Semantics and Structure in Statistical Translation*, pages 21–30. Association for Computational Linguistics.
- Hua Wu and Ming Zhou. 2003. Optimizing synonym extraction using monolingual and bilingual resources. In *Proceedings of the Second International Workshop on Paraphrasing - Volume 16, PARAPHRASE '03*, pages 72–79, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Shumin Wu, Jinho D. Choi, and Martha Palmer. 2010. Detecting Cross-lingual Semantic Similarity Using Parallel PropBanks. In *Proceedings of the 9th Conference of the Association for Machine Translation in the Americas, AMTA'10*, Denver, CO.
- Xian-mo Zeng. 2007. Semantic relationships between contextual synonyms. *US-China Education Review*, 4(9):33–37.