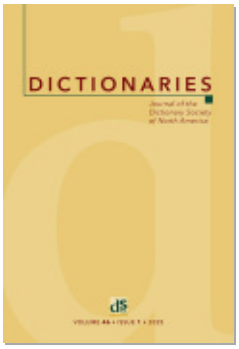
Linking an Event-type Ontology to Morphosyntax of the Predicate-Argument Structure



Zdeňka Urešová, Eva Fučíková, Cristina Fernández-Alcaina, Jan Hajič

Dictionaries: Journal of the Dictionary Society of North America,
Volume 46, Issue 1, 2025, pp. 207-227 (Article)

Published by Dictionary Society of North America
DOI: *https://doi.org/10.1353/dic.2025.a963331*

Linking an Event-type Ontology to Morphosyntax of the Predicate-Argument Structure

Zdeňka Urešová
Eva Fučíková
Cristina Fernández-Alcaina
Jan Hajič

Charles University

uresova@ufal.mff.cuni.cz
fucikova@ufal.mff.cuni.cz
alcaina@ufal.mff.cuni.cz
hajic@ufal.mff.cuni.cz

ABSTRACT

In the SynSemClass project, we are building a multilingual ontology for text annotation at the semantic (meaning) level. The SynSemClass ontology consists of entries (hereafter called *classes*) that represent one event type each, defined across multiple languages. Each class, representing an eventive type, contains a set of words (synonymous or semi-synonymous verbs) that express that event type, as described by its definition. Each verb is linked to both its syntactic properties and its occurrence in similar semantic lexicons for the particular language. Depending on the resource being referred to, the chain of links maps the particular class member to its sense as defined in that other resource. In addition, the semantic roles associated with the class are linked to the arguments (valency slots) as defined by the external resource, which in turn—in some cases—also contain morphosyntactic information relevant for each of the arguments. Such a mapping allows for the extraction of the correct case, adposition, word-order precedence and/or negation as well as other properties important for

the corresponding surface form. In this paper, we present the system of the linked data as present in SynSemClass, together with examples taken from the languages covered.

Keywords: lexical semantics, language resources, digital lexicons, meaning representation, valency, morphology, syntax

INTRODUCTION

This paper describes an event-type ontology, called SynSemClass, being built to investigate verbal synonymy in multilingual context and to relate semantic roles common to one synonym class to verb arguments (predicate-argument structure). In addition, this lexical resource is linked to existing resources with a similar aim: English WordNet, FrameNet, PropBank, VerbNet (SemLink), and valency lexicons for Czech, English, German and Spanish (PDT-Vallex, Vallex and EngVallex, E-VALBU, AnCora and other corpus-related lexicons). The SynSemClass ontology will have multiple uses: for meaning representation annotation of running textual corpora, where event types will be assigned to event instances described in the text; for event typology in knowledge bases; for comparative lexicographical work or creating highly structured dictionaries; as well as for other linguistic research uses, such as comparison to other similar resources in terms of structure, lexical coverage, paraphrasing or synonymy, or investigating the form of expression of the individual semantic roles for each event type and each lexical unit (in terms of syntax and morphology). Linking and mapping elements of SynSemClass entries to the Predicate-Argument Structure is the focus of this article.

THE SYNSEMCLASS ONTOLOGICAL LEXICON

The SynSemClass event-type ontology (Urešová et al. 2020, Urešová et al. 2023a, Urešová et al. 2023b) is being built as a richly structured and richly interlinked resource for use in meaning representation annotation and for various linguistic and lexicographic research purposes, as briefly mentioned in the Introduction section. Here, we describe its overall structure, the structure of each entry and the system of external links.

Overall structure. The overall structure is quite simple: the ontology consists of a set of entries (mostly verbs for now), each representing an eventive concept, such as RUNNING, THINKING, EATING or BEING LOCATED AT (states and events are not formally distinguished). The entries, however, are conceived as language-independent, defined by the cognitive content of the event type being represented. Each entry is called a *class*.[1]

There is also an assumed hierarchy of the classes, similarly to other lexical semantic resources. This hierarchy is currently being built). The hierarchy will link classes only as a whole (i.e., not individual class members). It should form a fully connected rooted tree, with each node (ideally) linked to one class. As an example of the hierarchy, the concept of flying will be linked (possibly through a chain, or path in the hierarchy tree, of more and more general event types) to the more general concept of motion.

Technically, the whole ontology consists of a number of files: the main structure, the hierarchy, and then one file for each language that is used to describe the entries.

The class (ontology entry) and its semantic roles. While the overall structure of the ontology is simple, as described in the previous section, the structure of each entry (class) is rich and complex.

With each class, there is a small set of semantic roles[2] that defines the relation of the given class to other entity types (as represented in some imaginary representation of text meaning), which, in a document referring to such concept, will be directly connected to it and described by that particular set of semantic roles. The set of roles is fixed for

[1]The concept of SynSemClass classes is in some aspects (multilinguality, synonymy etc.) not the same as the term *class* as defined for other resources, such as VerbNet, and it also differs from some theoretical concepts, such as Levin's classes (Levin 1993). However, if the hierarchy over the classes is taken into account, the upper levels of the hierarchy will correspond much more closely to the usual concept of "class" as defined for the existing resources.

[2]Similarly as for the use of the term *class*, the definition of the term *semantic role* in SynSemClass is not the same as in other lexical resources; given the focus on uniformity across languages and the cognitive aspect, perhaps it might be more fitting to call them *cognitive roles* or, from the perspective of their use in text, *semantic participants*; we often use quite specific "roles" like *Creator* and *Created* (e.g., for the class *Paint*), instead of *Agent* and *Theme*, which are used in other resources. However, we use the term *semantic role*, given that this is the term used previously when describing this resource.

any given class, regardless which word or expression might be used in an utterance to convey it. The morphosyntactic properties (usually language-dependent) are not ignored (see below the description of the relation between the semantic roles and valency or predicate-argument structure). The roles have their definitions, independent of the class to which they are linked; the same role can thus be assigned to different classes (keeping its definition), provided that the classes are "compatible" enough to allow for such sharing.

The presence of semantic roles at each class is nothing really new; it can be found in other well-known lexical resources as well, for example, in the form of "core arguments" in FrameNet (ref) or "obligatory arguments" in valency lexicons, such as VALLEX (Lopatková et al. 2022), PDT-Vallex (Urešová et al. 2021) and EngVallex (Cinková 2006, Cinková et al. 2021). However, what distinguishes the SynSemClass ontology from these resources is the requirement of strict language independence and the requirement that the roles with the same name use the same meaning and definition, even across classes. For example, the SynSemClass role *Source* is defined as 'Sb/sth from which sth originates or can be obtained'. As such it is used in 31 classes with 484 class members. By contrast, in FrameNet, some frame elements of the same name can have different meaning in different frames; for example, the FrameNet's *Source* is defined as 'the initial location of the Theme, before it changes location' (in **Removing** frame), or as 'the individual or environmental reservoir of a Disease that the Protagonist was exposed to' (in **Come_down_with** frame), or as 'the location that the Theme formerly occupied' (in Getting frame). Another example of one frame element having different meanings is the frame element called *Theme*, defined, for example, as 'The objects being carried' (in **Bringing** frame), or 'The object that moves' (in **Arriving** frame), or 'The object which is offered for transfer' (in **Offering** frame), or 'The objects being sent' (in **Sending** frame), or 'The item planted by the Agent' (in **Planting** frame).

Similar to SynSemClass, VerbNet uses also one definition for one role only; for example, VerbNet's *Source* is defined as 'The starting point (possibly metaphoric) of an action; it exists independently of the event'. However, VerbNet's 39 thematic roles are further specified with selectional restrictions that indicate the type of entity that usually fulfills that

role (Brown 2024). In contrast, SynSemClass contains 407 roles but no selectional restrictions.

Each class, in addition, contains *class members*, lexical units (consisting of a single word, such as *buy*, or a multiword expression [MWE], such as *look up* or *kick the bucket*, etc.) that can possibly be used as a form for expressing the eventive class in running text (in its surface form). There are, typically, multiple words (or MWEs, or a mix of single words and MWEs) in each class, for each language covered (the current version contains class members in English, German, Spanish, and Czech). It should be stressed that this set of class members is not necessarily complete; that is, there might be other words in any given language that can, in particular situations and contexts, express the same concept even if not included in the class. In a sense, the class members show only the *typical* expressions for the class's concept, but no strict division is being made between "norm" and "exploitation," as defined by Hanks (2013); if attested in data during the creation or maintenance of the entry, the class member (a word or MWE) is included in the class even if the context is not truly "canonical." The class members can be thus seen as a set of near-synonyms representing varying degrees of similarity, the use of which heavily relies on their context. In other words, the class members are more of a "hint" to human users (annotators, researchers using the resource, etc.), helping them to understand the class's concept, for which they can also rely on the class plain-text definition, the set of semantic roles and their definitions, the links to other resources, etc.

Each class member is technically identified by a unique ID (unique also across all the languages included in the ontology). Whenever possible, this ID points to (or is taken from) a syntactic resource which describes the valency or predicate-argument structure of that lexical unit: for each language, such a resource is always identified before the language is added to the ontology. This is important for two reasons:

- To be able to anchor the linking between the semantic roles and the surface form of morphosyntactic expression to an existing resource and annotation system;
- To be able to identify a particular sense of the class member in case of homonymy, since this is usually how the valency or predicate-argument resources are structured.

For the principles of the mapping, and how senses are handled in particular cases, see below ("Valency And Predicate-Argument Structure"). For an example, see Figure 1.

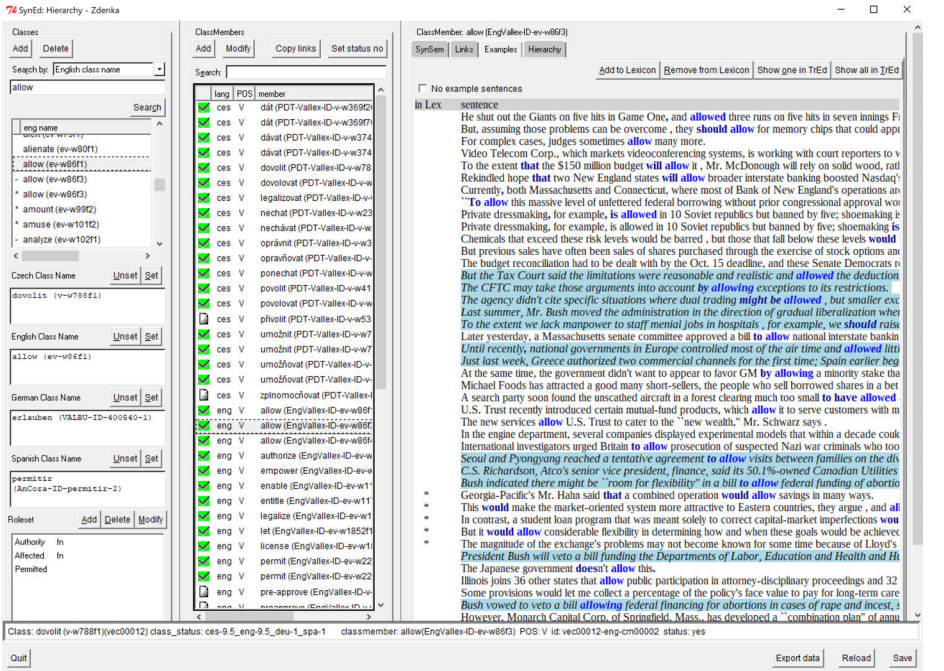


FIGURE 1 Example of a class (*allow*) in the viewer and editor SynEd, with examples

Links to external resources. On top of the valency and predicate-argument resources as described in the previous section (and in more detail below), each class member (i.e., not a class as a whole) is linked to the closest entry (or entries) in several similarly-focused resources that already exist. These links are thus also language-dependent, given that each class member is a lexical unit in a given language.

There, resources are linked mostly on a conceptual basis: if the sense of the class member (taking into account also the very concept of the class in which it appears) closely matches an entry in these resources, an appropriate link is added. Not surprisingly—in fact, quite often—this means that there are multiple links from one class member to different entries even in the same external resource, given the different structure, definitions and focus of these resources. We consider the existence of more links for one class member an advantage rather than a drawback: it allows comparison

of the resources in question, exemplifies the differences, and even finds inconsistencies and mistakes, often in a fast, automatic way.

The resources used for linking are as follows (excluding valency and predicate-argument lexicons—see below for their list and more details about their use in the role-form mappings).

For English,

- FrameNet (Baker et al. 1998), https://framenet.icsi.berkeley.edu, version framenet_v17-1
- WordNet (Fellbaum 1998), https://wordnetcode.princeton.edu/wn3.1.dict.tar.gz
- VerbNet (Kipper et al. 2006, Brown et al. 2022) https://github.com/cu-clear/verbnet/tree/master/verbnet3.4, also available from https://verbs.colorado.edu/verbnet/
- PropBank (Palmer et al. 2005, Pradhan et al. 2022) https://propbank.github.io/v3.4.0/frames/, also available from https://propbank.github.io/
- OntoNotes (Weischedel et al. 2013) https://clear.colorado.edu/compsem/index.php?page=lexicalresources&sub=ontonotes

For German,

- FrameNet des Deutsches (Ziem 2020), https://framenet-constructicon.hhu.de/framenet/frameindex
- Woxikon, https://synonyme.woxikon.de

For Czech,

- CzEngVallex (Hajič et al. 2016) https://lindat.mff.cuni.cz/services/CzEngVallex/ http://hdl.handle.net/11234/1-1512
- VALLEX 4.0 (Lopatková et al. 2022), http://hdl.handle.net/11234/1-3524

For Spanish,

- Spanish WordNet 3.0, integrated within the Multilingual Central Repository (MCR) (Gonzalez-Agirre et al. 2012), https://adimen.si.ehu.es/web/MCR
- Spanish Verbal SenSem Lexicon (Fernández et al. 2004), http://grial.edu.es/sensem/lexico/main

- ADESSE (García-Miguel et al. 2005), http://adesse.uvigo.es/
- Spanish FrameNet (Subirats 2009), https://sfn.spanishfn.org/SFN-reports.php

These resources differ widely in their coverage (Fučíková et al. 2024), focus (corpus annotation, general lexicon, etc.), stratum (some are more "semantic" than others), and availability of thorough description. The linking is thus not always perfect; despite that, we believe that the information contained in the links is very useful for the reasons mentioned above.

Current version of SynSemClass and quantitative description. The current version of SynSemClass is 5.0 (Urešová et al. 2023a). It is available for download at http://hdl.handle.net/11234/1-5230 (also available online for browsing at https://lindat.cz/services/SynSemClass50/ and for searching at https://lindat.mff.cuni.cz/services/SynSemClass-Search/?version=synsemclass5.0) (Petliak et al. 2023, Petliak et al. 2024). This version of the SynSemClass ontology has been used for the description and examples in the next main section of the article. See also Figure 2 for an example of an entry as shown in the SynSemClass browser.

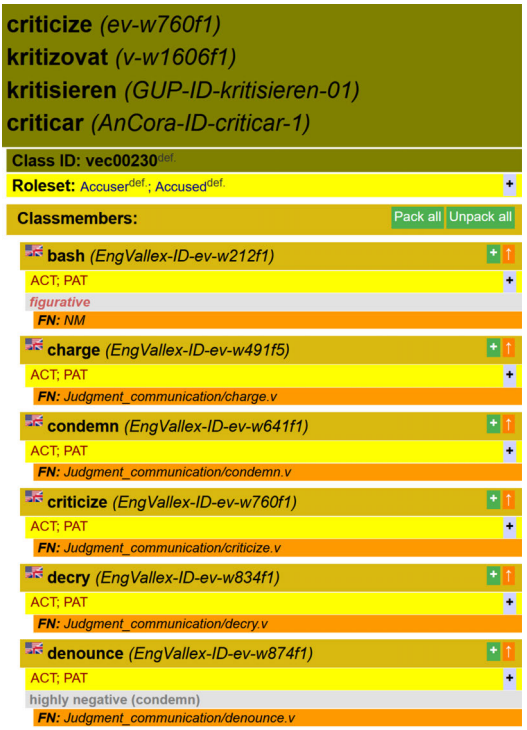


FIGURE 2 Example of the SynSemClass entry **criticize**

It has currently 1,546 classes and 15,790 lexical units in four languages (English, German, Spanish, and Czech, combined). For the detailed statistics on the newest version of the SynSemClass ontology, see http://hdl.handle.net/11234/1-5808.

VALENCY AND PREDICATE-ARGUMENT STRUCTURE

Resources used. For mapping the semantic roles associated with every class in SynSemClass with valency and predicate-argument structure we are using the following valency or predicate-argument resources:

- for English, EngVallex (Cinková et al. 2021) https://lindat.mff.cuni.cz/repository/xmlui/handle/11234/1-3526,
- for German, German Universal Propositions (Akbik et al. 2015) http://alanakbik.github.io/UniversalPropositions_German/index.html and also E-VALBU (Kubczak 2014) https://grammis.ids-mannheim.de/verbs,
- for Spanish, AnCora lexicon (Taulé et al. 2008) https://clic.ub.edu/corpus/en/ancoraverb_es,
- for Czech, PDT-Vallex (Urešová et al. 2014) http://hdl.handle.net/11858/00-097C-0000-0023-4338-F.

These resources serve two purposes: First, the entry in them identifies (in most cases) the sense of the lexical unit (mostly verb) and typically bears a unique ID for easy identification of the class member, and its sense in case of homonymy. Second, the entry (for an example, see Figure 3) in them contains the valency frame (or list of arguments) with the morphosyntactic description of each argument (valency slot) and its properties. These arguments (valency slots) serve as the elements to be connected to the semantic roles (see "Mapping semantic roles to arguments" below) of the class in which the verb (sense) resides.

FIGURE 3 EngVallex entry—**criticize** valency frame

If the particular sense matching the SynSemClass class member is not found in these resources, it has to be created and considered a "virtual" valency frame entry, as if it exists in one of the existing resources. The naming of the arguments is created in such a way that it cannot be confused with the existing entries, but at the same time, the structure of the arguments is created in such a way that it uses the same principles as the (or one of the) valency resources used for the particular language. Such entries are used as a last resort. In cases like German, where the coverage of a single resource may be insufficient, two resources are used to avoid over-reliance on "virtual" entries. For example, for the German verb *einräumen*, in the class **admit/přiznat se/gestehen**, a virtual valency entry has been created with three arguments, marked simply SA0 (corresponding with ACT), SA1 (corresponding with PAT), and SA2 (corresponding with ADDR), while the usual GUP arguments are labeled A0 (corresponding with ACT), A1 (corresponding with PAT), A2 (corresponding with ADDR), and the E-VALBU arguments would be VA0 (corresponding with ACT), VA1 (corresponding with PAT), and VA2 (corresponding with ADDR).

Mapping semantic roles to arguments. The mapping requires that each semantic role of the class, in which the class member (the verb, in this case) resides, is mapped to one of its arguments (valency slots) as found in the resource from which its (sense or valency frame) ID was taken.

For example, for the concept of 'CRITICIZING (somebody—someone)', in the class that represents it, there are two semantic roles assigned: 'the one who criticizes' (role: Accuser) and 'the one who is criticized' (role: Accused). In English, the EngVallex valency lexicon, which contains the correct sense of the verb (frame/sense 1), shows ACT (Actor) and PAT (Patient, or Theme) as two arguments for this sense (see Figure 3), making it natural to map Accuser to ACT and Accused to PAT (see Figure 4). In fact, all the English verbs from this class are in a similar situation: all have ACT and PAT listed as the two obligatory arguments in their valency frames (such as *charge* (frame/sense 5) or *condemn* (frame/sense 1)). For the other languages, the situation is equally simple: all the Czech class members (*kárat*, *kritizovat*, *hanobit*, *očerňovat*, *odsuzovat*, etc.) also have ACT and PAT as the obligatory arguments. In German, while the marking of the arguments is different (A0 and A1, for Argument 0 and Argument 1 as in PropBank), the situation is in principle the same and *kritisieren* has Accuser mapped to A0 and Accused to A1. We find the same case in Spanish, except that the arguments use a slightly different set of labels (arg0, arg1): *criticar*'s arg0 is mapped to Accuser and arg1 to Accused.

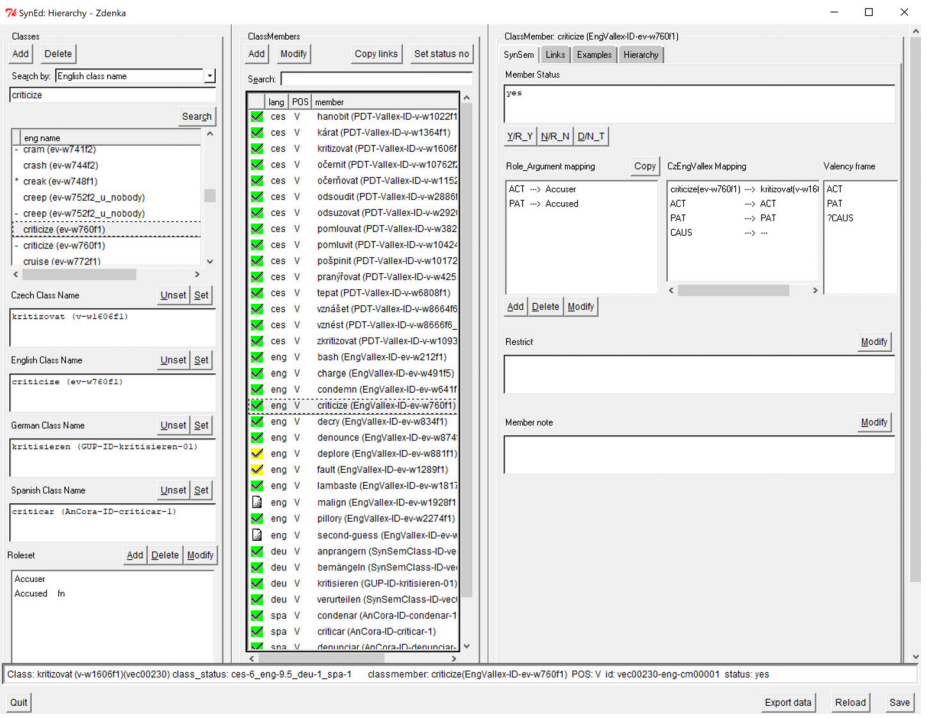


FIGURE 4 SynSemClass entry: **criticize** class, **criticize** member mapping

The situation however might be much more complex, and that is when the flexibility of the mapping is an advantage. For example, the English verbs in the class serve (as in *John served runny eggs and burnt toast to Mary*.) are associated with different sets of valency slots in the EngVallex lexicon: *serve* itself has only two actants (ACT, PAT) and a free modification ("non-core" or "optional" argument) labeled BEN is used for the person served (see Figure 5). For the verb *pass*, three obligatory (core) valency slots are used: ACT, PAT and DIR3 (see Figure 6, direction "to"). For the verb *present*, another member of the *serve* class, the usual ACT, PAT, and ADDR obligatory valency slots form its valency frame (Figure 7).

FIGURE 5 EngVallex entry—**serve** (sense 4) valency frame

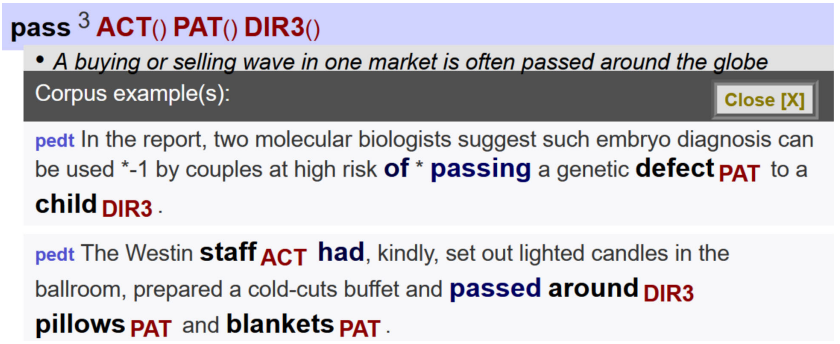


FIGURE 6 EngVallex entry—**pass** (sense 3) valency frame

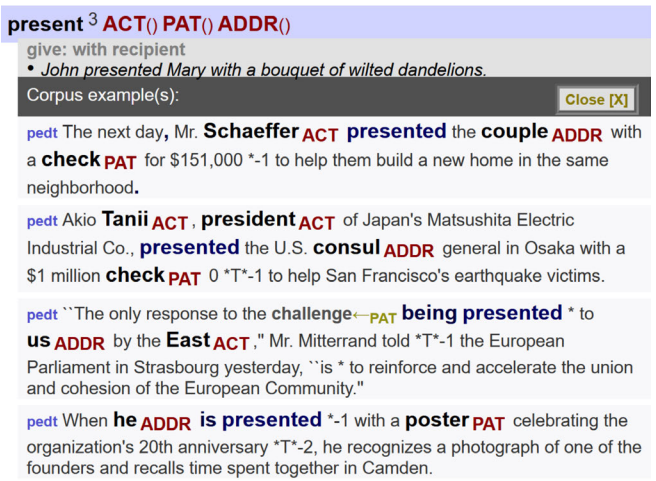


FIGURE 7 EngVallex entry—**present** (sense 3) valency frame

When mapping these class members' valency slots to the class's semantic roles, the BEN, DIR3 and ADDR are all mapped (at each of the verbs separately) to the semantic role of Recipient, since all are used to express it on the surface (see Figures 8, 9, and 10, respectively).

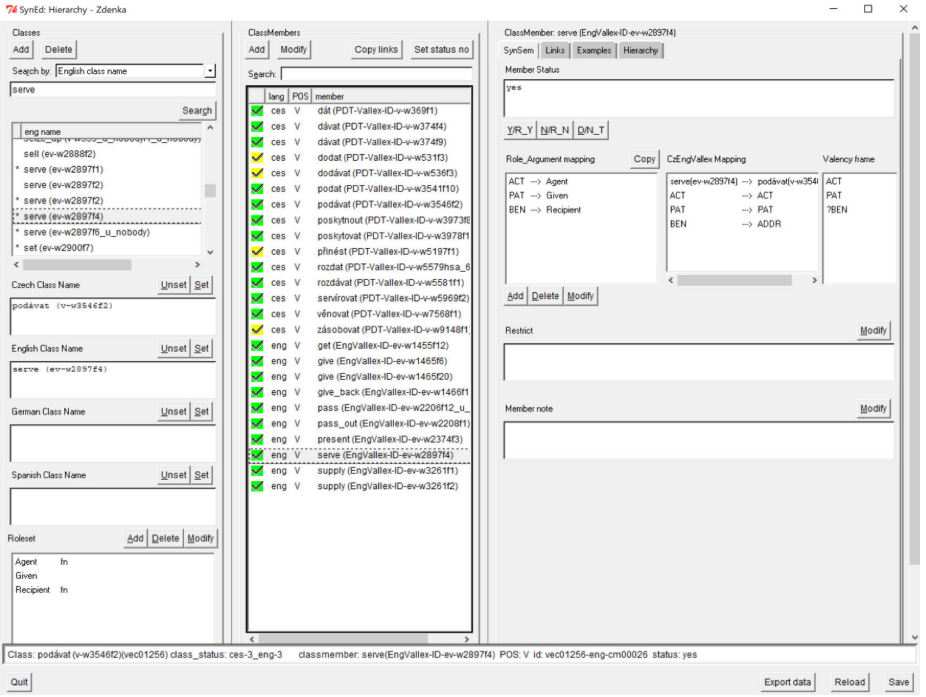


FIGURE 8 SynSemClass entry: **serve** class member mapping

In some cases, there might be a mismatch between the number of arguments (valency slots) of the class member and the set of semantic roles present in the class. While this should happen only rarely (overwriting the principle of semantic roles–arguments correspondence), it cannot be excluded, for two reasons: i) the principles of the valency theory used in the resources, from which the valency information is extracted, might vary, or ii) simply because the conceptual similarity "wins" in the sense that it describes the concept equally well as do other class members, except that one of the roles might not be present as an argument of the verb but (clearly) understood from the context. For example, in the class labeled as **help**, which is associated with three semantic roles (Helper, Party_benefited, and Goal) there are the following English verbs: *aid* and *help*, both with three arguments (mapped easily to the three semantic roles of the class), but also *help_out*, *support*,

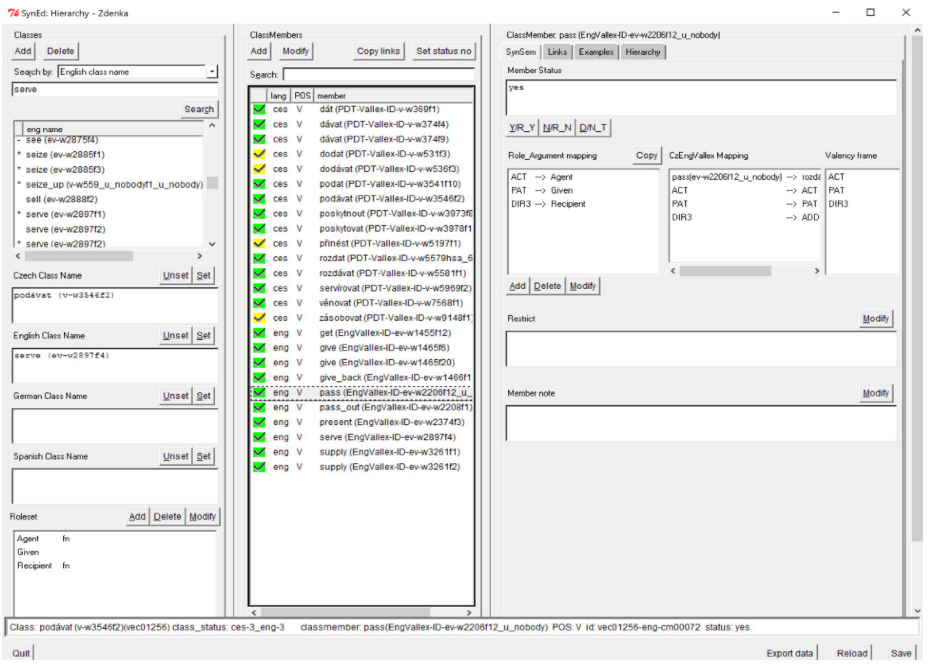


FIGURE 9 SynSemClass entry: **pass** class member mapping

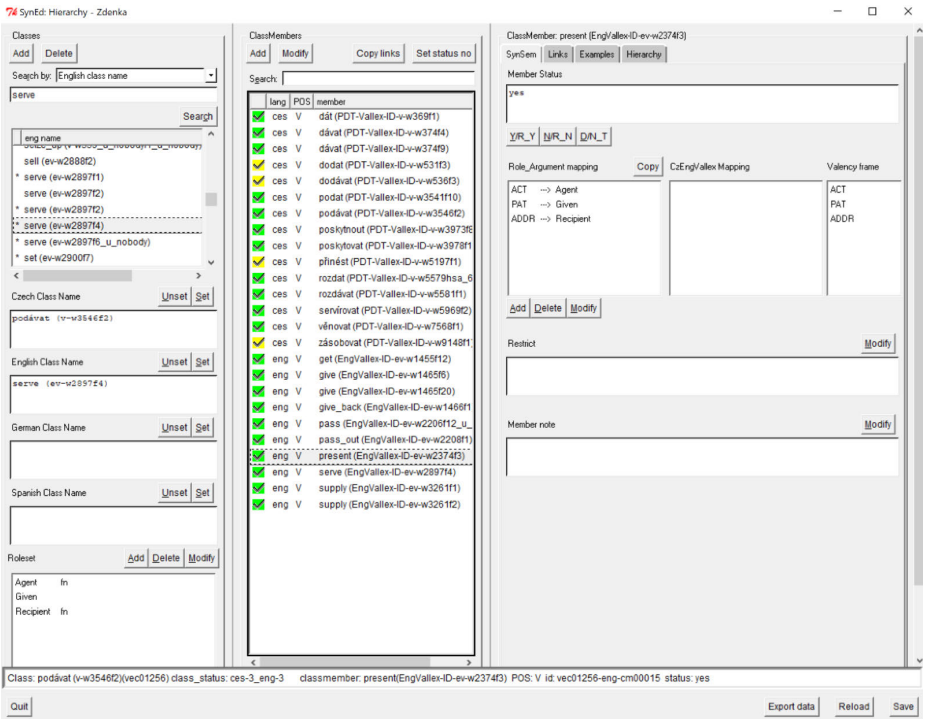


FIGURE 10 SynSemClass entry: **present** class member mapping

and *endorse*, which have (in EngVallex) only two arguments, but where the Party_benefited (or the Goal, depending of the alternation used) is assumed to be present in the surrounding text (in the same clause, or perhaps even in a different sentence), for example, in the sentence fragment *Sullivan.ACT will support Dr. Mason's ruling.PAT*, the role of Party_benefited is expressed by the possessive form *Dr. Mason's*, which does not fit any of the verb arguments.

These mappings, due to the explicit linking to the arguments, allow for extraction of any information from the valency lexicon(s) that is associated with the entry as a whole or with the individual arguments. For example, in the Czech PDT-Vallex lexicon, each valency slot contains—in addition to the label (ACT, PAT, EFF, etc.)—also a detailed list of possible surface realizations (forms allowed for the participant to take morphologically, if it appears at that position in the sentence). The simplest indication is that ACT usually appears in nominative and PAT in accusative; however, many other possibilities are listed too, given the verb to which the valency frame is associated. For example, the Czech verb *debatovat* 'to debate' can express ACT in nominative only but has a whole spectrum of possible surface forms for its PAT argument: typically, the preposition *o* 'about' with locative case is used, but a subordinate sentence attached through the conjunctions *že*, *zda*, *zdali* can be used too. In addition, instead of the preposition *o*, the multiword preposition *na téma* 'on the topic of' with accusative case can also be used, and given English influence on contemporary Czech, the accusative case alone (with no preposition) is now also tolerated as a correct morphosyntactic expression. The class in which *debatovat* appears (with also more general words (class members) that can mean the same in a particular context, such as 'talk about', 'chat', etc.) has many other Czech verbs: each of them has different means of PAT argument expression (for example, the colloquial *klábosit* does not allow— for PAT slot—any other form but *o* with locative case). The class member-based mapping then associates the semantic roles with the correct set of forms of expression for each member.

CONCLUSIONS AND FUTURE WORK

We have presented the way in which the SynSemClass event-type ontology maps its entries (more precisely, the verbs in its conceptual eventive classes) to existing valency and predicate-argument lexicons. It shows that the semantic roles associated with the whole class can map to different arguments, with different surface (morphosyntactic) realizations, and that the mapping also can generalize over different predicate-argument and/or predicate valency theories, for example between the PropBank-style argument definitions and the Functional Generative Description valency theory (Sgall et al. 1986), as reflected in the Prague Dependency Treebank resources (Hajič et al. 2020). The resource(s) are all freely available, and reachable through the links made accessible through the SynSemClass web project.[3]

In the future, on top of adding the hierarchy (which, however, will only minimally influence the role-argument mappings, and only in cases in which some class will be changed, split or merged with others), the mappings for new class members and new classes will be added and used mainly for checking the upcoming annotation projects for consistency between annotations on morphological, syntactic and semantic levels. At the same time, the mapping will serve as an empirical basis (given that all the entries contain examples from real texts) for linguistic studies on valency, predicate-argument structure, morphology and their relations to semantics, as well as for cross-linguistic studies at least for the languages currently available (and more, since some of the resources already being linked to are contained in the LLOD [Linguistic Linked Open Data[4]] database with additional connection to many more resources, including other languages).

ACKNOWLEDGMENTS

The work described herein has been supported by the project LUSyD, grant No. GX20-16819X of the EXPRO program of the Czech Science Foundation, and it has used resources and services of the LINDAT/CLARIAH-CZ Research Infrastructure, grant No. LM2023062 of the Ministry of Education, Youth and Sports of the Czech Republic.

[3]https://ufal.mff.cuni.cz/synsemclass
[4]https://linguistic-lod.org/

REFERENCES

Akbik, Alan, Laura Chiticariu, Marina Danilevsky, Yunyao Li, Shivakumar Vaithyanathan, and Huaiyu Zhu. 2015. Generating high quality Proposition Banks for multilingual semantic role labeling. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, 397–407. Beijing: Association for Computational Linguistics.

Baker, Collin F., 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*, 86–90. Stroudsburg, PA: Association for Computational Linguistics.

Brown, Susan Windisch. 2024. Extending VerbNet's verb-specific features to enhance selectional preferences of semantic roles. In *Proceedings of the Fifth International Workshop on Designing Meaning Representations @ LREC-COLING 2024*, 124–30, Torino: ELRA and ICCL.

Brown, Susan Windisch, Julia Bonn, Ghazaleh Kazeminejad, Annie Zaenen, James Pustejovsky, and Martha Palmer. 2022. Semantic Representations for NLP using VerbNet and the Generative Lexicon. *Frontiers in Artificial Intelligence* 5: 821697. doi: 10.3389/frai.2022.821697

Cinková, Silvie. 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)*, 2170–75. Genova, Italy: European Language Resources Association (ELRA).

Cinková, Silvie, Eva Fučíková, Jana Šindlerová, and Jan Hajič. 2021. *EngVallex: English Valency Lexicon 2.0*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, http://hdl.handle.net/11234/1-3526.

Fellbaum, Christiane. 1998. *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

Fernández, Ana, Gloria Vázquez, and Irene Castellón. 2004. Sensem: base de datos verbal del español. In *IX Ibero-American Workshop on Artificial Intelligence (IBERAMIA 2004)*, 155–63. Puebla de los Ángeles, Mexico.

Fučíková, Eva, Cristina Fernández-Alcaina, Jan Hajič, and Zdeňka Urešová. 2024. Textual coverage of eventive entries in lexical semantic resources. In *Proceedings of the 2024 Joint International Conference on Computational Linguistics,*

Language Resources and Evaluation (LREC-COLING 2024), 15835–41. Torino: ELRA and ICCL.

García-Miguel, José M., Lourdes Costas, and Susana Martínez. 2005. Diátesis verbales y esquemas construccionales. Verbos, clases semánticas y esquemas sintáctico-semánticos en el proyecto ADESSE. In *Entre semántica léxica, teoría del léxico y sintaxis*, edited by Gerd Wotjak and Juan Cuartero Otal, 373–84. Berlin: Peter Lang.

Gonzalez-Agirre, Aitor, Egoitz Laparra, and German Rigau. 2012. Multilingual central repository version 3.0. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*, 2525–29. Istanbul: European Language Resources Association (ELRA).

Hajič Jan, Eva Fučíková, Jana Šindlerová, and Zdeňka Urešová. 2016. Verb argument pairing in Czech–English Parallel Treebank. In *GLOBALEX 2016: Lexicographic Resources for Human Language Technology, at LREC 2016*, 16–23.

Hajič, Jan, et al. 2020. *Prague Dependency Treebank–Consolidated 1.0 (PDT-C 1.0)*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, http://hdl.handle.net/11234/1-3185.

Hanks, Patrick. 2013. *Lexical Analysis: Norms and Exploitations*. Cambridge, MA: MIT Press.

Kipper, Karin, Anna Korhonen, Neville Ryant, and Martha Palmer. 2006. Extending VerbNet with novel verb classes. In *Proceedings of the 5th International Conference on Language Resources and Evaluation*, 1027–32. Genova, Italy: European Language Resources Association (ELRA).

Kubczak, Jacqueline. 2014. *Valenzwörterbuch E-VALBU*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, http://hdl.handle.net/11372/LRT-1162.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. University of Chicago Press.

Lopatková, Markéta, Václava Kettnerová, Jiří Mírovský, Anna Vernerová, Eduard Bejček, and Zdeněk Žabokrtský. 2022. *VALLEX 4.5*. LINDAT/CLARIAH-CZ digital Library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, Prague, http://hdl.handle.net/11234/1-4756.

Palmer, Martha, Daniel Gildea, and Paul Kingsbury. 2005. The Proposition Bank: An annotated corpus of semantic roles. *Computational Linguistics*, 31(1): 71–106.

Petliak, Nataliia, Jan Hajič, Zdeňka Urešová, and Eva Fučíková. 2023. *SynSemClass Search Tool*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, http://hdl.handle.net/11234/1-5281.

Petliak, Nataliia, Cristina Fernández-Alcaina, Eva Fučíková, Jan Hajič, and Zdeňka Urešová. 2024. Search tool for An Event-Type Ontology. In *Proceedings of the Twentieth Workshop on Interoperable Semantic Annotation*, 66–70. Torino: ELRA and ICCL.

Pradhan, Sameer, Julia Bonn, Skatje Myers, Kathryn Conger, Timothy O'Gorman, James Gung, Kristin Wright-Bettner, Martha Palmer. 2022. PropBank comes of age—larger, smarter, and more diverse. In *Proceedings of* the 11th Joint Conference on Lexical and Computational Semantics, 278–88.

Sgall, Petr, Eva Hajičová, and Jarmila Panevová. 1986. *The Meaning of the Sentence in its Semantic and Pragmatic Aspects*. Dordrecht: Reidel.

Subirats, Carlos. 2009. Spanish FrameNet: A frame-semantic analysis of the Spanish lexicon. In *Multilingual FrameNets in Computational Lexicography: Methods and Applications*, edited by Hans C. Boas, 135–62. Berlin, New York: De Gruyter Mouton.

Taulé, Mariona, M. Antònia Martí, and Marta Recasens. 2008. AnCora: Multilevel annotated corpora for Catalan and Spanish. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08)*, 96–101. Marrakech: European Language Resources Association (ELRA).

Urešová, Zdeňka, Jan Štěpánek, Jan Hajič, Jarmila Panevová, and Marie Mikulová. 2014. *PDT-Vallex: Czech Valency lexicon linked to treebanks*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, http://hdl.handle.net/11858/00-097C-0000-0023-4338-F.

Urešová, Zdeňka, Eva Fučíková, Eva Hajičová, and Jan Hajič. 2020. SynSemClass linked lexicon: Mapping synonymy between languages. In *Proceedings of the 2020 Globalex Workshop on Linked Lexicography (LREC 2020)*, 10–19. Marseille: European Language Resources Association (ELRA).

Urešová, Zdeňka, Alevtina Bémová, Eva Fučíková, Jan Hajič, Veronika Kolářová, Marie Mikulová, Petr Pajas, Jarmila Panevová, and Jan Štěpánek. 2021. *PDT-Vallex: Valenční slovník češtiny propojený s korpusy 4.0*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics

(ÚFAL), Faculty of Mathematics and Physics, Charles University, http://hdl.handle.net/11234/1-3499.

Urešová, Zdeňka, et al. 2023a. *SynSemClass 5.0*. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University, http://hdl.handle.net/11234/1-5230.

Urešová, Zdeňka, Cristina Fernández-Alcaina, Eva Fučíková, and Jan Hajič. 2023b. SynSemClass Czech and English Annotation Guidelines. Technical report no. 2023/73. *The ÚFAL Technical Report Series*. https://ufal.mff.cuni.cz/techrep/tr73.pdf.

Weischedel, Ralph, Martha Palmer, Mitchell Marcus, Eduard Hovy, Sameer Pradhan, Lance Ramshaw, Nianwen Xue, Ann Taylor, Jeff Kaufman, Michelle Franchini, Mohammed El-Bachouti, Robert Belvin, and Ann Houston. 2013. *OntoNotes Release 5.0*. LDC2013T19. Philadelphia: Linguistic Data Consortium. https://doi.org/10.35111/xmhb-2b84.

Ziem, Alexander. 2020. Wortbedeutungen als Frames: ein Rahmenmodell zur Analyse lexikalischer Bedeutungen. In *Semantiktheorien II: Analysen von Wort- und Satzbedeutungen im Vergleich*, edited by Sven Staffeldt and Jörg Hagemann, 27–56. Tübingen: Stauffenburg.