

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

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

Institute of Formal and Applied Linguistics
Computer Science School
Faculty of Mathematics and Physics
Charles University, Prague, Czech Republic

Supported by LUSyD (GAČR, GX20-16819X), UMR (MŠMT ČR, LUAUS23283) and Mellon Foundation (G-1901-06505), using resources of LINDAT/CLARIAH-CZ RI (MŠMT ČR, LM2018101, LM2023062).

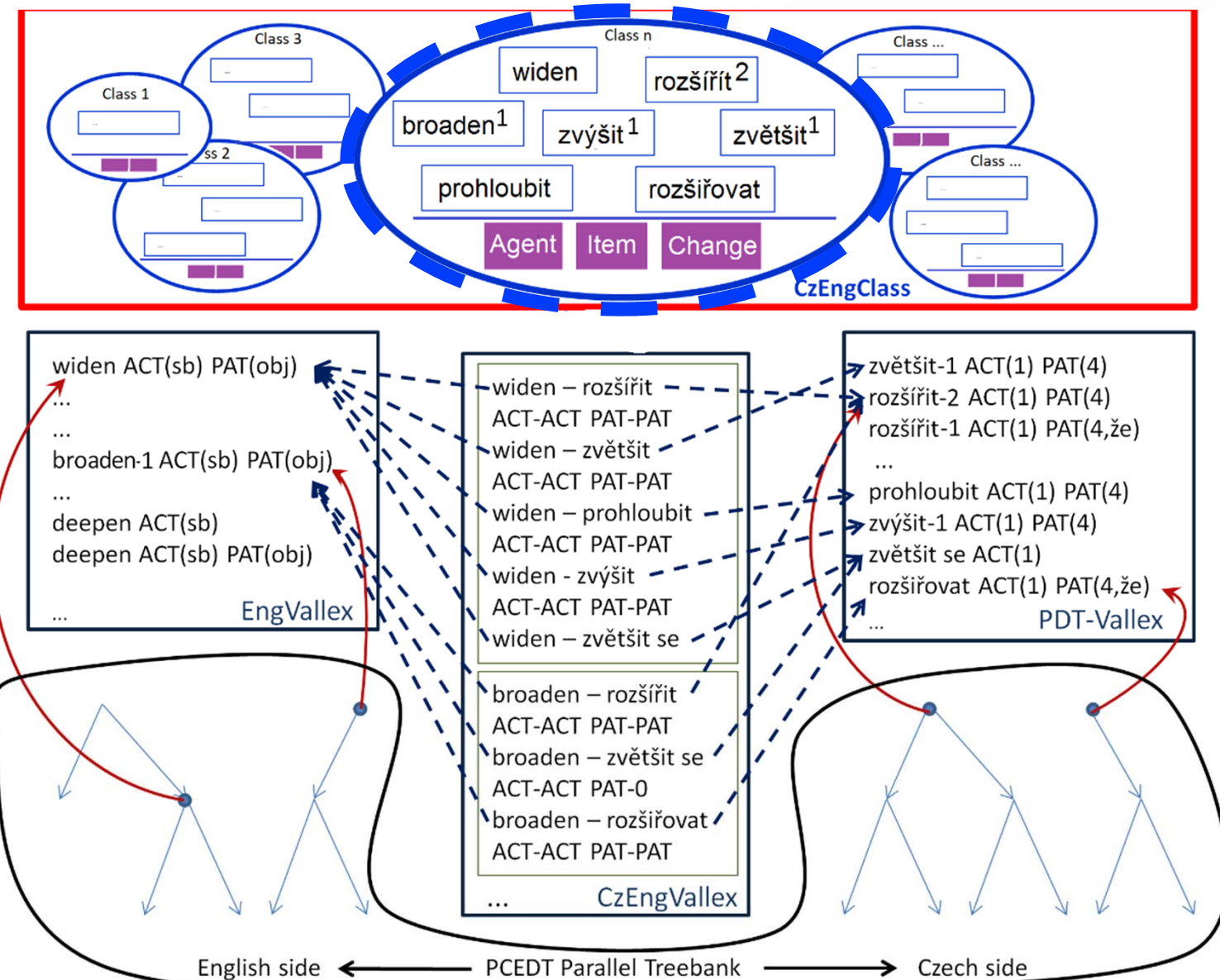


Outline

- SynSemClass as an event-type ontology
 - For annotation of events and event types in texts
- SynSemClass in a nutshell
 - Classes (concepts), Semantic Roles
- Links to syntactic behaviour of lexical units
 - Links to external lexico-semantic resources
- Conclusions and Future Work

SynSemClass

- Event-type ontology (a.k.a. verb synonym lexicon)
 - ~900 classes, Czech – English
 - grouping verb senses with similar meaning - **Class members**
 - Common set of **Semantic Roles** per class (Roleset)
- Links to external resources, examples
 - FrameNet, WordNet, VerbNet, ...
- Semantic roles
 - Mapped to syntactic arguments
- SynEd – Annotation editor
- Web browsing (+ free for download)
 - <https://lindat.cz/services/SynSemClass40>



Editor for ontology maintenance (SynEd)

- Using SynEd
 - Special purpose editor for SynSemClass
- Language

SynEd: SynSemClass LEXICON (GE)

Classes

Add Delete

Search by: Class lemma

vinit

Search

lemma

vinit (v-w7617f1)

- vést (v-w7583f5)

* vést (v-w7583f8)

* vést si (v-w7586f3_ZU)

- vidět (v-w7612f5)

* vidět (v-w7612f16_ZU)

- vidět (v-w7612f17_ZU)

* vinit (v-w7617f1)

* viset (v-w7622f1)

vládnout (v-w7641f1)

* vlastnit (v-w7650f1)

vlévat se (v-w11270f1)

* vnútit (v-w7678f1)

* vracet se (v-w7703f2)

Set

English Class Name

accuse (ev-w31f1)

Roleset

Add Delete Modify

Accuser

Addressee fn

Cause fn

Note

Modify

ClassMembers

Add Modify Copy links

Search:

member	
blame (EngVallex-ID-ev-w272f1)	En
blame (EngVallex-ID-ev-w272f2)	En
charge (EngVallex-ID-ev-w491f2)	En
dávat (PDT-Vallex-ID-v-w374146_ZU)	Cz
nařknout (PDT-Vallex-ID-v-w2199f1)	Cz
obvinit (PDT-Vallex-ID-v-w2584hsa-602)	Cz
obviňovat (PDT-Vallex-ID-v-w2586f1)	Cz
obžalovat (PDT-Vallex-ID-v-w2597f1)	Cz
vyčítat (PDT-Vallex-ID-v-w7825f1)	Cz
žalovat (PDT-Vallex-ID-v-w10020f1)	Cz
charge (EngVallex-ID-ev-w491f3)	En
countersue (EngVallex-ID-ev-w728f2)	En
impeach (EngVallex-ID-ev-w1652f1)	En
indict (EngVallex-ID-ev-w1684f1)	En
sue (EngVallex-ID-ev-w3248f1)	En
zažalovat (PDT-Vallex-ID-v-w9324f1)	Cz
žalovat (PDT-Vallex-ID-v-w10020f3)	Cz
sue (EngVallex-ID-ev-w3248f2)	En
anklagen (GUP-ID-anklagen-01)	De
anlasten (GUP-ID-anlasten-01)	De
anschuldigen (SYNSEMCLASS-ID-ect00347cm0145)	De
belangen (GUP-ID-belangen-01)	De
beschuldigen (GUP-ID-beschuldigen-01)	De
bezüchtigen (GUP-ID-bezüchtigen-01)	De
geben (SYNSEMCLASS-ID-vec00347cm0145)	De
machen (SYNSEMCLASS-ID-vec00347cm0145)	De
schleichen (VALBU-ID-400863-2)	De
unterstellen (GUP-ID-unterstellen-01)	De

ALL YES NO NOT_TOUCHED

RATHER_YES RATHER_NO Deleted

ClassMember: unterstellen (GUP-ID-unterstellen-01)

SynSem Links Examples

FrameNet Des Deutschen

Urteil 843

VALBU

NO MAPPING

GUP

unterstellen/unterstellen.01

ParaCrawl German

unterstellen accuse

Export data Reload Save

class: vinit (v-w7617f1)(vec00347) class_status: 4_de1 classmember: unterstellen(GUP-ID-unterstellen-01) id: vec00347cm00145 status: yes

Concepts vs. words

- SynSemClass structure
 - Language-independent “core” – classes (concepts), roles
 - Language-dependent form of expression / definitions
 - For humans to understand the class’ concept
- Creation process
 - Core: deciding on concepts, roles, definitions
 - Language-dependent: adding words (lexical units)
 - Including **information on syntactic behavior**
 - External links and examples

Arguments and roles 1

- Concept (class; example: “**promise**”)
 - Fixed list of characteristic roles (same across languages)
- Class members ~ verbs in different languages
 - Each has some argument structure
 - English: **promise**
 - PDT-Vallex: ACT PAT ADDR
 - PropBank: Arg0-PAG, Arg1-GOL, Arg2-PPT
 - Czech: **slíbit** (promise)
 - PDT-Vallex: ACT PAT ADDR
 - Class roles: Communicator, Commitment, Recipient
- Mapping:

Role	Argument (cs, en)
Communicator	ACT (Arg0)
Commitment	PAT (Arg2)
Recipient	ADDR (Arg1)

Arguments and roles 2

- Not always simple 1:1
 - Alternation
 - Different arguments for different verbs („looser“ synonymy)
- Alternation example
 - English: **erase** ACT PAT ORIG
 - PropBank: Arg0-PAG Arg1-PPT Arg2-DIR Arg3-MNR
 - Czech: **vymazat**¹ (erase), PDT-Vallex: ACT PAT DIR1 (... file from computer)
 - **vymazat**² (erase), PDT-Vallex: ACT PAT (... computer)
 - Class roles: Protagonist_Remover, Removed, Place

- Mapping:

Role	Argument (en)	Argument (cs, vymazat ¹)	Argument (cs, vymazat ²)
Protagonist	ACT (Arg0)	ACT	ACT
Removed	PAT (Arg1)	PAT	
Place	ORIG (Arg2)	ORIG	PAT



Arguments and roles 3

- Mapping across arguments and across languages
 - English:
 - **link up** ACT PAT ADDR [MEANS]
 - PropBank: Arg0-PAG Arg1-PPT Arg2-PPT Arg3-MNR
 - **ally** ACT PAT[refl.] EFF
 - PropBank: Arg0-PAG Arg1-PPT Arg2-COM
 - Czech:
 - **splynout** (merge), PDT-Vallex: ACT PAT[with] EFF[into]
 - Class roles: Participant_1 Participant_2 Whole
- Mapping:

Role	Argument (en, link up)	Argument (en, ally [refl.])	Argument (cs, splynout)
Participant_1	ACT (Arg1)	ACT (Arg0)	ACT
Participant_2	PAT (Arg2)	PAT (Arg1)	PAT
Whole	ADDR (??)	EFF (Arg2)	EFF



SynSemClass lexicon: browse (v3.5)

© 2021 Univerzita Karlova v Praze, ÚFAL MFF UK (Charles University in Prague, Czech Republic)

Please see also the [Contextually-based synonymy and valency of verbs in a bilingual setting](#) [main research page](#) for additional resources and documentation.

- ☒ **FrameNet des Deutschen**
- ☒ **Woxikon**
- ☒ **E-VALBU**
- ☒ **GUP**
- ☒ **Examples**
- ☒ **Check all**
- ☐ **Uncheck all**

Select languages:

- ☐ **English**
- ☐ **Czech**
- ☒ **German**

Show classes with German name starting with...

A B D E G L M V Ü

Classes in SynSemClass with German name starting with the letter 'b':

- [Beschleunigen \(GUP-ID-beschleunigen-01\)](#)
- [Beschuldigen \(GUP-ID-beschuldigen-01\)](#)
- [Bewerben \(SYNSEMCLASS-ID-vec00276cm00047\)](#)

urychlit (v-w7267f1)
accelerate (ev-w18f2)
beschleunigen (GUP-ID-beschleunigen-01)

Class ID: vec00337

Roleset: Agent; Difference; Item_changed; Value_final; Value_initial

FN: Causation; Cause_change_of_position_on_a_scale; Cause_to_start; Subjective_influence

Classmembers:

ankurbeln (SYNSEMCLASS-ID-vec00337cm00052)

SA0; SA2(um); SA1; SA4(auf, zu); SA3(von)

GFN: NM

WOX: ankurbeln#5; ankurbeln#9

EVA: NM

GUP: NM

NO EXAMPLES

beschleunigen (GUP-ID-beschleunigen-01)

A0; A2(um); A1; A4(auf, zu); A3(von)

GFN: NM

WOX: beschleunigen#7

EVA: NM

GUP: beschleunigen/beschleunigen.01

Know-how, hohe Kapazitäten und die Kombination verschiedener Verfahren beschleunigen die Projektabwicklung.
Speed Knowledge, large capacities and the combination of various processes accelerate project execution.

Access to SynSemClass

The class (multilingual concept)

Class' common roleset

Roleset to Arg mapping

Class members

Conclusions and Future Work

- Results and conclusions
 - 6,105 English, 6,037 Czech, 533 German, ~500 Spanish verbs
 - Total of ~900 classes, freely available
 - Rich contents (mappings, links to other lexical resources)
- Future Work
 - Extend SynSemClass with further languages (Korean)
 - Link to more lexical resources via LLOD
 - Comparison with automatic synonym discovery methods
 - Explore possibility of semi-automatic checks, extensions
 - LAW paper (ACL 2023): using LLM to predict class for unknown word
 - Annotate data (PDT/TR, UMR, ...)

Thank you!



References

- Zdeňka Urešová, Karolina Zaczynska, Peter Bourgonje, Eva Fučíková, Georg Rehm, Jan Hajič. 2022. Making a Semnatic Event-type Ontology Multilingual. LREC 2022, Marseille, France (this presentation).
- Petr Sgall, Eva Hajičová, and Jarmila Panevová. 1986. *The meaning of the sentence in its semantic and pragmatic aspects*. D. Reidel, Dordrecht.
- 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, and Jana Šindlerová. 2016. *CzEngVallex: a bilingual Czech-English valency lexicon*. The Prague Bulletin of Mathematical Linguistics, 105:17–50.
- 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. Defining verbal synonyms: between syntax and semantics. In Dag Haug, Stephan Oepen, Lilja Ovrelid, Marie Candito, and Jan Hajič, editors, *Proceedings of the 17th International Workshop on Treebanks and Linguistic Theories (TLT 2018)* (Pub. No. 155), pages 75–90, Linköping, Sweden. Universitetet i Oslo, Linköping University Electronic Press.
- Zdeňka Urešová, Eva Fučíková, Eva Hajičová, and Jan Hajič. 2018c. Synonymy in Bilingual Context: The SynSemClass Lexicon. In *Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018*, Santa Fe, New Mexico, USA, August 20-26, 2018, pages 2456–2469.