

A data-based analysis of the effect of prefixation on the syntactic-semantic characteristics of verbs in Czech

Hana Hledíková, hana.hledikova@ufal.mff.cuni.cz

5th International Workshop on Resources and Tools for Derivational Morphology

University of Fribourg, Switzerland

4–5 September, 2025



Institute of Formal and Applied Linguistics
Faculty of Mathematics and Physics
Charles University



Structure of the presentation

1) Introduction

Prefixation in Czech

Effect of prefixation on valency

2) Research objective

3) Data

Morphological resources

Valency lexicons

SynSemClass

4) Results

5) Discussion

Prefixation in Czech

Prefixation in Czech

- **Prefixation** = a new word is created via addition of a word-formation prefix
- Very **productive in Slavic languages** (e.g. Körtvélyessy 2016)
- Connected to **grammatical aspect**
- Meaning on different levels of concreteness, **polyfunctionality**

syp-a-t > **vy**-*syp-a-t*
pour-THEME-INF out-pour-THEME-INF
'pour (impf.)' 'pour out (pf.)'

ps-á-t > **vy**-*ps-a-t*
write-THEME-INF out-write-THEME-INF
'write (impf.)' 'write down out of sth / issue, call (e.g. elections) (pf.)'

syp-a-t > **na**-*syp-a-t*
pour-THEME-INF on-pour-THEME-INF
'pour (impf.)' 'pour on (pf.)'

ps-á-t > **na**-*ps-a-t*
write-THEME-INF on-write-THEME-INF
'write (impf.)' 'write (pf.)'

- **Resultative analysis** of prefixation
 - “the event leads (...) to some impact, effect, consequence, goal – in general to a **result**, which was not implied by the base verb” (Uher 1987: 44)
 - “prefixes introduce a **causal relation** btw. **the state brought about** by the PP and the event introduced by the verbal root” (Biskup 2019: 8)

Effect of prefixation on valency

- Prefixes can **change valency**

(Jirsová 1979, Bémová 1981, Uher 1987, Svenonius 2004, Romanova 2006, Ramchand 2008, Biskup 2019)

- Resultative analysis predicts **tendency to add a valency slot**

- Direct object (patient)

zářít > ***o-zářít celé poschodí***
'shine' 'around-shine (= illuminate) the whole staircase'

- Prepositional phrase

- Location

psát poznámky > ***vy-psat poznámky z encyklopedie***
'write notes' 'out-write notes from the encyclopedia'

- Result

tlačit na vojáky > ***do-tlačit vojáky k zásahu***
'push, pressure the soldiers' 'in-push the soldiers into action'

- Indirect object (addressee)

psát recept > ***přede-psat pacientovi recept***
'write a prescription' 'prescribe the patient a prescription'

Research objective

Research objectives

Analyze the **patterns of changes in valency** that happen in **Czech prefixation** using a **large data sample** collected using a combination of **morphological** and **valency data resources**

- **Morphological data resources:** pairs of 'prefixed verb – base verb' (instead of selected examples)
- **Lexicon of verb senses:** mapping of individual senses of the prefixed and base verbs
- **Valency lexicon:** valency characteristics for the senses of the prefixed and base verbs
- Use the sample to **analyze the most frequent patterns of change**
- Verify the **expectations** from previous theoretical literature:
 - Resultative nature of the prefix → Addition of valency slots (patient, location, result, addressee)

Data

- Extraction of pairs of **base verb** – **prefixed verb** in Czech
 - **List of verbs** from SYN2000 (Čermák et al. 2000) → **19 000 verb lemmas**
 - Annotated for **morphemic structure and family membership** using available morphological resources (Slavíčková 1975, Vidra et al. 2021) combined with manual checking
→ **8 361 pairs of prefixed verbs and their bases**
- Extraction of **valency frames** for the verbs
 - **Valency lexicons: Vallex** (Lopatková et al. 2022), **PDT-Vallex** (Urešová et al. 2024)
 - Functional Generative Description (e.g. Panevová 1974, 1975), functor labels
 - Arguments:
 1. **ACT** = actor (***Peter** broke the vase. **The vase** broke.*)
 2. **PAT** = patient (*Peter broke **the vase**.*)
 3. **ADDR** = addressee (*Peter gave **her** the letter.*)

ORIG = origin (*Peter received a letter **from his mother**.*)

EFF = effect (*Peter turned his room **into a gym**.*)
 - Adjuncts (**MANN** = manner, **TWHEN** = time, **DIR1** = source, **DIR2** = path, **DIR3** = goal, ...)
 - Issue: Vallex does **not** contain **mapping between senses** of derivationally related verbs

Data: SynSemClass (version 5.1)

- Synonymy lexicon of verbs (Urešová et al. 2024)
- Groups **verbs' senses into classes** – semantic equivalence (synonymy / near synonymy)

hrát 'to play' – 3 senses:

class	vec00611
definition	'A Performer plays a Role.'
roles	Performer, Role

class	vec00823
definition	'A Performer plays or transmits Music'
roles	Performer, Music

class	vec00415
definition	'A Competitor 1 plays against a Competitor 2 in a competition for a Targeted entity.'
roles	Competitor 1 and 2, Competition, Targeted

- Linked to PDT-Vallex

word (sense)	valency dictionary entry	valency frame
<i>hrát</i> (vec00611)	<i>PDT-Vallex-ID-v-w1122f2</i>	ACT, PAT
<i>hrát</i> (vec00823)	<i>PDT-Vallex-ID-v-w1122f4</i>	ACT, PAT
<i>hrát</i> (vec00415)	<i>PDT-Vallex-ID-v-w1122f1</i>	ACT, PAT, ADDR, EFF

Data: sense-to-sense mapping

- Vallex does **not** contain **mapping between senses of derivationally related verbs**
- SynSemClass helps:

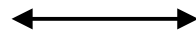
1) When prefixed verb and base verb in **the same class**

hrát 'to play (impf.)'

classes	roles
vec00611 'play a role'	<i>Performer, Role</i>
vec00823 'play music'	<i>Performer, Music</i>
vec00415 'play in a game/competition'	<i>Competitor_1, Competitor_2, Competition, Targeted</i>

za-hrát 'to play (pf.)'

classes	roles
vec00611 'play a role'	<i>Performer, Role</i>
vec00823 'play music'	<i>Performer, Music</i>



Data: sense-to-sense mapping

- Vallex does **not** contain **mapping between senses of derivationally related verbs**
- SynSemClass helps:

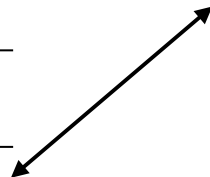
2) When there is an **overlap in semantic roles**

hrát 'to play'

classes	roles
vec00611 'play a role'	<i>Performer, Role</i>
vec00823 'play music'	<i>Performer, Music</i>
vec00415 'play in a game/competition'	<i>Competitor_1, Competitor_2, Competition, Targeted</i>

pro-hrát 'to lose'

classes	roles
vec00275 'to lose in a game/competition'	<i>Loser, Competition, Winner</i>



- Automatic pre-annotation based on SynSemClass → manual annotation

Data: sense-to-sense mapping

- Final sample:

lemma pairs	both lemmas in SynSemClass	lemmas with at least one sense-sense mapping	sense-sense pairs
8 361	1 076	630	869

- Change in valency for each sense-sense pair

hrát 'play in a game/competition' (vec00425): ACT PAT ADDR EFF
prohrát 'lose' (vec00275): ACT PAT ADDR } -EFF

spát 'sleep' (vec00518): ACT
prospat 'sleep through something' (vec01279): ACT PAT } +PAT

Results

Changes in valency frames

[illegible]

red = additions, yellow = deletions

Changes in valency frames: PAT

[illegible]

red = additions, yellow = deletions

Changes in valency frames: PAT

<i>zář-i-t</i> shine-THEME-INF 'to shine'	>	<i>o-zář-i-t</i> around-shine-THEME-INF 'to shine onto something, illuminate'	<i>něc-o</i> something-ACC.SG PAT
---	---	---	--

<i>síl-i-t</i> strong-THEME-INF 'to become stronger'	>	<i>po-síl-i-t</i> on-strong-THEME-INF 'to strengthen, make something stronger'	<i>něc-o</i> something-ACC.SG PAT
--	---	--	--

<i>mysl-e-t</i> think-THEME-INF 'to think'	>	<i>vy-mysl-e-t</i> out-think-THEME-INF 'to think up, invent something'	<i>něc-o</i> something-ACC.SG PAT
--	---	--	--

<i>sp-á-t</i> sleep-THEME-INF 'to sleep'	>	<i>pro-sp-a-t</i> through-sleep-THEME-INF 'to sleep through something'	<i>něc-o</i> something-ACC.SG PAT
--	---	--	--

Changes in valency frames: ADDR/ORIG/EFF

[illegible]

red = additions, yellow = deletions

Changes in valency frames: -ADDR/EFF

hlás-i-t *něk-omu* *něc-o*
 call-THEME-INF somebody-DAT.SG something-ACC.SG
 ADDR **PAT**
 ‘to report something to somebody’

>

vy-hlás-i-t *něc-o*
 out-call-THEME-INF something-ACC.SG
 PAT
 ‘to announce something’

dar-ova-t *něk-omu* *něc-o*
 gift-THEME-INF somebody-DAT.SG something-ACC.SG
 ADDR **PAT**
 ‘to gift something to somebody’

>

ob-dar-ova-t *někoho*
 around-call-THEME-INF something-ACC.SG
 PAT
 ‘to give somebody a gift’

soud-i-t *o něč-em* *že je to špatné*
 judge-THEME-INF about something-LOC.SG that it is bad
 PAT **EFF**
 ‘to judge something to be bad’

>

od-soud-i-t *něc-o*
 away-judge-THEME-INF something-ACC.SG
 PAT
 ‘to condemn something’

hodnot-i-t *něc-o* *jako důležité*
 value-THEME-INF something-ACC.SG as important
 PAT **EFF**
 ‘to evaluate something as important’

>

nad-hodnot-i-t *něc-o*
 over-value-THEME-INF something-ACC.SG
 PAT
 ‘to overvalue, overestimate something’

Changes in valency frames: DIR

[illegible]

red = additions, yellow = deletions

Changes in valency frames: DIR

<i>téc-i</i>	>	<i>při-téc-i</i>	<i>do jezera</i>
flow-INF		in-flow-INF	into lake-GEN.SG
			DIR3
'to flow'		'to in-flow into the lake'	

<i>let-ě-t</i>	>	<i>od-letět</i>	<i>z Čech</i>
fly-THHEME-INF		away-fly-THHEME-INF	from Bohemia
			DIR1
'to fly'		'to fly away from Bohemia'	

<i>let-ě-t</i>	>	<i>od-letět</i>	<i>do Čech</i>
fly-THHEME-INF		away-fly-THHEME-INF	from Bohemia
			DIR3
'to fly'		'to fly away to Bohemia'	

<i>bod-nou-t</i>	<i>tyč</i>	<i>do zem-ě</i>	>	<i>pro-bod-nou-t</i>	<i>zem</i>
stab-THHEME-INF	pole	into ground-GEN.SG		through-stab-THHEME-INF	gound
	PAT	DIR3			PAT
'to stick a pole into the ground'				'to pierce the ground'	

Discussion

Discussion

- Changes in **PAT**: confirmed that almost always **addition**
- Changes in the **third valency slot ADDR/ORIG/EFF**
 - Alongside expected additions, we also find a significant number of **deletions**
- The resultative nature of the prefix does not necessarily lead to an addition of a participant (and therefore a valency slot), but **can actually lead to deletions of valency slots**
- Sometimes the prefix leads to a **change in focus** of attention
 - Resulting speech act more salient at the expense of the recipient

hlásit někomu něco 'to report sth to sb' > *vy-hlásit něco* 'to announce sth'

- Recipient more salient at the expense of the theme

darovat někomu něco 'to gift sth to sb' > *ob-darovat někoho* 'to give sb a gift'

- Landmark more salient at the expense of the theme

bodnout tyč do země 'to stick a pole into the ground' > *pro-bodnout zem* 'to pierce the ground'

Discussion

Limitations:

- Coverage, particular theoretical framework – captures some patterns, not others
- Changes in terms of what is considered a part of the valency frame or not – **binary distinction**
- **In reality a matter of degree (frequency of occurrence)**

bodnout tyč do země ‘to stick a pole into the ground’ – *probodnout zem (tyčí)* ‘to pierce the ground (**with a pole**)’

- High level of generalization in defining the valency slots – some changes not captured

<i>pít pivo</i>	>	<i>pro-pít své peníze</i>
PAT		PAT
‘drink a beer’		‘drink away your money’

Future:

- Closer look at the verb pairs and their semantics: **larger classes of verbs** and their behaviour
- SynSemClass also includes a meaning hierarchy
 - do semantically similar verbs have similar valency changes during prefixation?

Summary

- Analysis of valency changes in prefixation in Czech
- Using combination of word-formation and valency data resources
- **Confirmed** expectations from theoretical literature: **addition of valency slots**
- Data also **revealed deletions**, especially in the third argument slot
- Resultativity of the prefix **interacts with valency in varied ways**, not only connected to additions of participants, but also **changes in focus**
- **Patterns found in a larger data sample** are able to reveal these interactions

References

- Bémová, A. (1981). Slovesná prefixace z hlediska intence. *Slovo a slovesnost*, 42(2): 143-148.
- Biskup, P. (2019). Prepositions, Case and Verbal Prefixes. *The Case of Slavic*. John Benjamins.
- Čermák, F. et al. (2000). SYN2000: Žánrově vyvážený korpus psané češtiny. Ústav Českého národního korpusu FF UK.
- Jirsová, A. (1979). Prefixace sloves a slovesná vazba. *Naše řeč*, 62(1):1–7.
- Körtvélyessy, L. (2016). Word-formation in Slavic languages. *Poznan Studies in Contemporary Linguistics*, 52(3):455–501.
- Lopatková, L. et. al (2022) VALLEX 4.5. <http://hdl.handle.net/11234/1-4756>.
- Panevová, J. (1974). On verbal frames in functional generative description. Part I. *The Prague Bulletin of Mathematical Linguistics* 22:3–40.
- Panevová, J. (1975). On verbal frames in functional generative description. Part II. *The Prague Bulletin of Mathematical Linguistics* 23:17–52.
- Ramchand, G. (2008). Verb meaning and the lexicon: a first-phase syntax. Cambridge University Press, Cambridge, UK.
- Romanova, E. (2006). Constructing Perfectivity in Russian. Doctoral thesis, University of Tromsø.
- Slavíčková, E. (1975). Retrográdní morfemický slovník češtiny. Academia, Praha.
- Svenonius, P. (2004). Slavic prefixes inside and outside VP. In *Nordlyd* 32.2, 205–253. CASTL, Tromsø.
- Uher, F. (1987). Slovesné předpony. Univerzita J. E. Purkyně, Brno.
- Urešová, Z. et al. (2023). 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, Z. et al. (2024). PDT-Vallex: Czech valency lexicon linked to treebanks 4.5. 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/11234/1-5814>.
- Vidra, J. et al. (2021). DeriNet 2.1. 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/11234/1-5814>.

Acknowledgements



This contribution was supported by the Charles University, project GA UK No. 246723, and the Charles University Research Centre program No. 24/SSH/009.