NOMVallex: A Valency Lexicon of Czech Nouns and Adjectives

Veronika Kolářová, Anna Vernerová

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

Supported by the Czech Science Foundation under the project 22-20927S.
Outline

• Czech monolingual ‘VALLEX’ lexicons
• NomVallex 2.0 lexicon
  – structure of the lexicon
  – derivational types
  – description of a lexicon entry
• Relationship between valency behavior of base words and their derivatives
  – links from derived lexical units to their base lexical units
  – comparison of their valency
  – systemic and non-systemic valency behavior
Czech monolingual ‘VALLEX’ lexicons

• VALency LEXicons
• Valency theory of the Functional Generative Description
  – Panevová (1974-75, 1998, 2000), Piťha (1982)
Czech monolingual ‘VALLEX’ lexicons

• **VALency LEXicons**

• Valency theory of the Functional Generative Description
  – Panevová (1974-75, 1998, 2000), Piťha (1982)

- strictly related to the data of the Prague Dependency Treebank
- only the word senses encountered in the corpus
- only captures their valency frames

- based on the SYN series of CNC
- all senses of the words included
- multiple syntactic phenomena related to valency
- complex searches and comparisons

- e.g., diatheses, reciprocity

- e.g., non-systemic valency behavior

Video Presentation

LREC 2022, Marseille, June 20-25
Czech monolingual ‘VALLEX’ lexicons

- **VALency LEXicons**
- Valency theory of the Functional Generative Description
  - Panevová (1974-75, 1998, 2000), Piťha (1982)

PDT-
- strictly related to the data of
  the Prague Dependency Treebank
- only the word senses encountered in the corpus
- only captures their valency frames

VALLEX
- e.g., diatheses, reciprocity
- based on the SYN series of CNC
- all senses of the words included
- multiple syntactic phenomena related to valency
- complex searches and comparisons

NOM
- e.g., non-systemic valency behavior
### Part-of-speech category

| Part-of-speech category | Derivational category | Lexical units | Lexemes |
|------------------------|-----------------------|---------------|---------|
| **Nouns**              | deverbial             | 505           | 248     |
|                        | deadjectival          | 247           | 158     |
| **Adjectives**         | deverbial             | 179           | 116     |
|                        | denominal             | 28            | 14      |
|                        | deadjectival          | 6             | 6       |
|                        | primary               | 62            | 28      |
| **Total**              |                       | **1027**      | **570** |
Stem nouns, e.g., žádání ‘asking’ (containing a theme suffix)
Root nouns, e.g., žádost ‘request’ (not containing a theme suffix)

Deverbal adjectives – deverbal derivatives with adjectival inflection:
- 9 types, derived for example from transgressives or participles
- cf. derivatives of the verb podezírat ‘to suspect’, e.g., suspected or suspicious

podezírající -- podezřelý -- podezíraný podezíratelný podezíravý -- podezřelý
**NomVallex**: An example of a lexicon entry

**Lexeme**: an abstract unit associating lexical forms with their lexical units

**Lexical unit (LU)**: word sense

**nadšení** `pf` *‘delight, joy’*

---

**ACT** 2, *pos* PAT 7, nad+7, z+2, že, cont
derivedFrom blu-v-nadchnout-2, blu-v-nadchnout-se-1

**PDT-vallex** v-w1991f1

**synon** velká radost

**examplerich**

**ACT+PAT** Nenušila nás, protože její **ACT nadšení** minulostí **PAT** bylo opravdu úžasné. • Vstřebal okouzlení svého otce přírodou, **nadšení** své matky **ACT** Grace umění **PAT** a každodenní zvyk svého dědečka Erneste „Abby” Halla vyprávět příběhy i jeho náboženské zánícení. • Fojtík by však raději viděl spolupráci s jinými ekonomikami a hejtmanovo **ACT nadšení** z Číny **PAT** nesdílí. •

... shortened ...

**ACT** S tak bezvýhradným **nadšením** publikace **ACT** i kritiky **ACT** jako tehdy u tohoto představení se setkáme jen zřídka.

**PAT** V **nadšení** nad příchodem **PAT** automobilky město i stát pozapomněly na řadu důležitých věcí. • Takové **nadšení** z vítězství **PAT** nepamatují. • Zrovna neomdléváš **nadšením**, že jsem **PAT** tady. •

Do spáťanské pokladny se sypaly peníze, rostlo **nadšení**, jak moderní fotbal tým hraje **PAT**.

---

**class** psych state

**semcategory** abstraktní výsledek děje

---

[https://ufal.mff.cuni.cz/nomvallex](https://ufal.mff.cuni.cz/nomvallex)
Lexeme: an abstract unit associating lexical forms with their lexical units
Lexical unit (LU): word sense

**nadšení** pf  ‘delight, joy’

[blu-n-nadšení-1]

- **ACT** 2, pos **PAT** 7, nad+7, z+2, že, cont
derivedFrom blu-v-nadchnout-2, blu-v-nadchnout-se-1  **ACT** 4 → 2, pos **PAT** 1 → 7 • že, cont, nad+7, z+2, inf, pro+4

dt-vvallex v-w1991f1

- **synon** velká radost

- **examplerich**  **ACT+PAT** Nenudila nás, protože její **ACT nadšení** minulostí **PAT** bylo opravdu úžasné. • Vstřebal okouzlění svého otce přírodou, **nadšení** své matky **ACT** Grace umění **PAT** a každodenní zvyk svého dědečka Ernesta „Abby“ Halla vyprávět příběhy i jeho náboženské zanícení. • Fojtík by však raději viděl spoluprací s jinými ekonomikami a hejtmanovo **ACT nadšení** z Číny **PAT** nesdílí. •

  ... shortened ...

  **ACT** S tak bezvýhradným **nadšením** publika **ACT** i kritiky **ACT** jako tehdy u tohoto představení se setkáme jen zřídká.

  **PAT** V **nadšení** nad příchodem **PAT** automobilky město i stát pozapomněly na řadu důležitých věcí. • Takové **nadšení** z větězství **PAT** nepamatují. • Zrovna neomdléváš **nadšení**, že jsem **PAT** tady. •

  Do spartanské pokladny se sypaly peníze, rostlo **nadšení**, jak moderní fotbal tým hraje **PAT**.

- **class** psych state
- **semcategory** abstraktní výsledek děje

[https://ufal.mff.cuni.cz/nomvallex](https://ufal.mff.cuni.cz/nomvallex)
**Lexeme:** an abstract unit associating lexical forms with their lexical units

**Lexical unit (LU):** word sense

nadšení *pf* ‘delight, joy’

**Valency frame:**
- actants or free modifications
- morphemic forms
- obligatoriness

**Combinations of actants**

**Corpus examples:**
- CNC SYNv8
- Araneum Bohemicum Maximum

https://ufal.mff.cuni.cz/nomvallex
Relationship between valency behavior of base words and their derivatives

• Examination of the extent to which nouns and adjectives share valency properties with their base words
  – argument inheritance hypotheses (Booij 2007)

• NomVallex
  – links from derived lexical units to their base lexical units
    • in case of verbs, to base verbal lexical units contained in VALLEX
  – automatic comparison of their valency frames
    • the number and type of valency slots
    • morphemic forms
  – the difference in valency is specified in the ‘valdiff’ attribute and set in color
  – systemic (regular) and non-systemic (irregular) valency behavior
    • Kolářová et al. 2019, 2021
**NomVallex: An example of a lexicon entry**

**Lexeme:** an abstract unit associating lexical forms with their lexical units

**Lexical unit (LU):** word sense

**nadšení** pf  ‘delight, joy’

**Link(s) to the base verbal LU(s) in VALLEX**

---

**Lexeme:** an abstract unit associating lexical forms with their lexical units

**Lexical unit (LU):** word sense

**difference in valency compared to the valency frame of the base verbal LU**

---

**valency frame:**

**ACT** 2, pos PAT 7, nad+7, z+2, že, cont
derivedFrom **blu-v-nadchout-2, blu-v-nadchout-se-1**
pdt-vallex **v-w1991f1**

**synon:** velká radost

**examplerich**

**ACT**+**PAT** Nenudila nás, protože její **ACT** nadšení minulostí **PAT** bylo opravdu úžasné. • Vstřebal okouzlení svého otce přírodou, nadšení své matky **ACT** Grace umění **PAT** a každodenní zvyk svého dědečka Ernesta „Abby“ Halla vyprávět příběhy i jeho náboženské zanícení. • Fojtík by však raději viděl spolupráci s jinými ekonomiky a hejtmanovo **ACT** nadšení z Číny **PAT** nesdílí. •

**ACT** S tak bezvýhradným nadšením publika **ACT** i kritiky **ACT** jako tehdy u tohoto představení se setkáme jen zřídka.

**PAT** V nadšení nad příchodem **PAT** automobilky město i stát pozapomněly na řadu důležitých věcí.

• Takové nadšení z vítězství **PAT** nepamatují. • Zrovna neomdléváš nadšením, že jsem **PAT** tady. •

Do spartánské pokladny se sypaly peníze, rostlo nadšení, jak moderní fotbal tým hraje **PAT**.

---

**class:** psych state

**semcategory:** abstraktní výsledek děje

---

[https://ufal.mff.cuni.cz/nomvallex](https://ufal.mff.cuni.cz/nomvallex)
Deverbal adjectives
Denominal adjectives
Deadjectival adjectives
Primary adjectives
Deadjectival nouns
Deverbal nouns

Interlinked lexical units

$\text{NomVallex} \rightarrow (\text{NomVallex}) \rightarrow \text{VALLEX}$

- žádanost ‘demand’ $\rightarrow$ žádaný-3 ‘in demand’ $\rightarrow$ žádat-1 ‘to demand’
- žádní-2 ‘asking’ / žádost-1 ‘request’ $\rightarrow$ žádat-2 ‘to ask’

Vallex
žádat-1
‘to demand’
žádat-2
‘to ask’

Deverbal adjectives
Denominal adjectives
Deadjectival adjectives
Primary adjectives
Deadjectival nouns
Deverbal nouns

Video Presentation
LREC 2022, Marseille, June 20-25
Deverbal adjectives
Denominal adjectives
Deadjectival adjectives
Primary adjectives
Deadjectival nouns
Deverbal nouns

VALLEX

- žádanost ‘demand’ > žádaný-3 ‘in demand’ > žádat-1 ‘to demand’
- žádání-2 ‘asking’ / žádost-1 ‘request’ > žádat-2 ‘to ask’

Diagram:
- žádání-2 ‘asking’
- žádost-1 ‘request’
- žádanost ‘demand’
- žádaný-3 ‘in demand’
- žádat-1 ‘to demand’
- žádat-2 ‘to ask’

Video Presentation
LREC 2022, Marseille, June 20-25
Comparison of valency properties:
Stem vs. root deverbal nouns

Verb:

žádat \textit{impf} žádávat \textit{iter} \textit{‘to ask’}

[blu-v-žádat-2]
‘to ask the employer. ADDR for a job. PAT’
žádat zaměstnavatele \textit{AK} ADDR o práci. PAT

Stem noun:

žádání (si) \textit{impf} \textit{‘asking’}

[blu-n-žádání-si-2]
‘asking the employer. ADDR for a job. PAT’
žádání zaměstnavatele \textit{Gen} ADDR o práci. PAT

Root noun:

žádost \textit{no-aspect} \textit{‘request’}

[blu-v-žádat-2]
‘request to-the-employer. ADDR for a job. PAT’
žádost zaměstnavateli \textit{Dat} ADDR o práci. PAT

Video Presentation
LREC 2022, Marseille, June 20-25
15
Comparison of valency properties:
Stem vs. root deverbal nouns

Verb:

**Stem noun:**

**Root noun:**
Comparison of valency properties: Stem vs. root deverbal nouns

Verb:

| Stem noun: | Root noun: |
|------------|------------|
| žádat impf žádávat iter | žádost no-aspect |
| ‘to ask’ | ‘request’ |

| žádat zaměstnavatele | žádost zaměstnavateli |
|----------------------|-----------------------|
| [blu-v-žádat-2] | [blu-v-žádat-2] |
| ‘to ask the employer. ADDR for a job. PAT’ | ‘request to-the-employer. ADDR for a job. PAT’ |

| žádání (si) impf | žádost (si) no-aspect |
|------------------|-----------------------|
| ‘asking’ | ‘request’ |

| žádání zaměstnavatele | žádost zaměstnavateli |
|----------------------|-----------------------|
| [blu-n-žádání-si-2] | [blu-v-žádat-2] |
| ‘asking the employer. ADDR for a job. PAT’ | ‘request to-the-employer. ADDR for a job. PAT’ |

| žádání zaměstnavatele | žádost zaměstnavateli |
|----------------------|-----------------------|
| derivedFrom blu-v-žádat-2 ACT | derivedFrom blu-v-žádat-2 ACT |
| ADDR o+4, aby, at', zda | ADDR o+4, inf, aby, at', zda, že |
| 2, 7, pos | 2, pos |
| PAT o+4, aby, at', zda | PAT o+4, aby, at', zda, inf, že |
| 1, 2, pos | 1, 2, pos |

Video Presentation

LREC 2022, Marseille, June 20-25
Comparison of valency properties:
Stem vs. root deverbal nouns

Verb:

**Stem noun:**

žádání (si) **impf**  ‘asking’

**Root noun:**

žádost **no-aspect**  ‘request’

---

Video Presentation  LREC 2022, Marseille, June 20-25
Comparison of valency properties: Stem vs. root deverbal nouns

Verb:

‘to ask’

Stem noun:

‘asking’

Root noun:

‘request’

VERB

[blu-v-žádat-2]

[blu-n-žádání-si-2]

[blu-v-žádat-2]
Comparison of valency properties: Stem vs. root deverbal nouns

Verb:

| Stem noun: | Root noun: |
|------------|------------|
| žádat impf žádávat iter ‘to ask’ | žádost no-aspect ‘request’ |
| žádat zaměstnavatele ADDR o práci PAT | žádost zaměstnavateli ADDR o práci PAT |

Stem noun:

| žádání (si) impf | žádání zaměstnavatele Gen ADDR o práci PAT |
|------------------|------------------------------------------|
| žádání zaměstnavatele Gen ADDR o práci PAT | žádost zaměstnavateli Dat ADDR o práci PAT |
| žádání zaměstnavatele Gen ADDR o práci PAT | žádost zaměstnavateli Dat ADDR o práci PAT |

Root noun:

| žádat impf žádávat iter | žádat zaměstnavatele ADDR o práci PAT |
|--------------------------|-------------------------------------|
| žádat zaměstnavatele ADDR o práci PAT | žádost zaměstnavateli ADDR o práci PAT |

Verbs:

| Verb | Stem noun | Root noun |
|------|-----------|-----------|
| žádat | žádat | žádat |
| žádávat | żádání | žádost |
| žádat zaměstnavatele | žádání zaměstnavatele | žádost zaměstnavateli |
Conclusions

• **NomVallex**
  – a manually annotated valency lexicon of Czech nouns and adjectives
  – comparison of valency of derivationally related lexical units

• Systemic and non-systemic valency behavior of various derivational types of Czech nominals
  – differences in valency between deverbal nouns and their base verbs and between deverbal adjectives and their base verbs
  – non-systemic valency behavior is mostly manifested by non-systemic forms of actants
References

• Booij, G. (2007). *The Grammar of Words*. Oxford: Oxford University Press.

• Kolářová, V., Vernerová, A., and Klímová, J. (2021). Systemic and Non-systemic Valency Behavior of Czech Deverbal Adjectives. *Jazykovědný časopis*, 72(2): 371–382.

• Kolářová, V., Vernerová, A., and Verner, J. (2019). Non-systemic valency behavior of Czech deverbal nouns based on the NomVallex lexicon. *Jazykovědný časopis*, 70(2): 424–433.

• Kolářová, V., Vernerová, A., and Klímová, J. (2022). NomVallex 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/11234/1-4663](http://hdl.handle.net/11234/1-4663), online search: [https://ufal.mff.cuni.cz/nomvallex](https://ufal.mff.cuni.cz/nomvallex).

• Lopatková et al. (2020). *VALLEX 4.0*, 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/11234/1-3524](http://hdl.handle.net/11234/1-3524).

• Panevová, J. (1980). *Formy a funkce ve stavbě české věty*. Praha: Academia, 222 p.

• Panevová, J. (1998). Ještě k teorii valence. *Slovo a slovesnost*, 59: 1–14.

• Panevová, J. (2000). Poznámky k valenci podstatných jmen. In Z. Hladká & P. Karlík (Eds.), *Čeština – univerzália a specifika*, 2, Brno: Vydavatelství MU, pp. 173–180.

• Piťha, P. (1982). K otázce valence u adjektiv. *Slovo a slovesnost*, 43, pages 113–118.

• Urešová a kol. (2021): PDT-Vallex. Valenční slovník češtiny propojený s korpusy 4.0. Praha: UFAL MFF UK.