

---

Veronika Kolářová and Jiří Mírovský

## LOOKING FOR SENSE IN NONSENSE

### Valency of Negative Forms of Nouns and Adjectives in the NomVallex Lexicon

**Abstract** NomVallex is a manually annotated valency lexicon of Czech nouns and adjectives that enables research into various language phenomena related to valency, including the comparison of valency properties of affirmative and negative forms of words. This paper presents new developments in the way the lexicon facilitates research into word-level negation, explaining the reasoning behind the proposed lexicographic treatment. Differentiating between direct negation and lexicalized negation, we discuss whether or not the negative forms of words should be listed separately in a valency lexicon. We argue that, while lexicalized negation has to be assigned an individual entry, direct negation of nouns and adjectives should be treated within an entry for affirmative forms. Considering various aspects of word-level negation, including the employment of negative forms of nouns and adjectives in light verb constructions or phrasemes, we describe negation-related attributes applied to the data of the lexicon. As a case study, the facilities provided by the lexicon are used to illustrate use and distribution of negative forms of Czech deadjectival nouns, and their valency properties are compared to those of the corresponding affirmative forms.

**Keywords** negation; negative form; affirmative form; nouns; adjectives; valency lexicon

## 1. Introduction

Negation can be expressed using various grammatical processes, especially sentential negation, phrasal negation and word-level negation (Curiel, 2015). Word-level (lexical) negation, e.g., *independence*, affects different word classes in different ways (Pavlovič, 2015), and thus may attract lexicographers' attention.

Kováříková et al. (2012) discuss whether or not the negative forms of words should feature as entries in a monolingual dictionary. Based on corpus data, the authors investigate whether negation of Czech verbs, nouns, adjectives and adverbs can be considered a grammatical category (Kováříková, 2011), and they analyze the coverage of negative forms (N-forms) and affirmative forms (A-forms) of verbs, adjectives and adverbs in different frequency levels. They conclude that, while N-forms of verbs should not be assigned an individual entry in the dictionary, N-forms of nouns, adjectives and adverbs should be listed separately (the two latter provided that they are frequent enough).

In specialized dictionaries, for instance, in valency lexicons (for English, see Herbst et al., 2004; for Czech, see Hajič et al., 2003, Svozilová et al., 2005), N-forms of nouns and adjectives are usually accorded a separate entry, too.

In this paper, we deal with the treatment of negation in NomVallex, a valency lexicon of Czech nouns and adjectives (Kolářová & Vernerová, 2022, Section 4). The aims of the study are twofold: (i) theoretical research into various aspects of lexical negation with respect to the meaning and valency of nouns and adjectives; (ii) description of the lexicographic treatment of negation in NomVallex. First, differentiating between direct negation and lexicalized negation, we outline the status of word-level negation in Czech word formation (Section 2). Further, we specify aspects that should be considered when treating negation in a valency lexicon (Section 3). Then we introduce the NomVallex lexicon (Section 4.1) and describe how negation is treated in the lexicon data (Section 4.2). Finally, as a case study, preliminary results of the annotation are illustrated through the case of negative forms of Czech deadjectival nouns and their valency (Section 5).

## 2. Word-Level Negation and Word Formation

From the word-formation point of view, word-level negation can be seen as the combination of a negative formant (negator) and an affirmative base (Pavlovič, 2015). European languages usually express word-level negation by affixation (Zimmer, 1964, Curiel, 2015, van Son et al. 2016). In Slavic languages, the universal exponent of negation is the morpheme *ne-* (Pavlovič, 2015; e.g., Cz. *nezávislost* ‘independence’), used with autosemantic word classes, i.e., nouns, adjectives, verbs and adverbs.

Focusing on Czech, when the negative prefix *ne-* is combined with verbs, it typically only denies the affirmative content, i.e., the original predication, without specifying new features (Lotko, 1975; cf. *záviset* ‘depend’ vs. *nezáviset* ‘not to depend’); therefore, N-forms of verbs are not treated in any specific way in Czech lexicons. In contrast, with nouns, adjectives and adverbs, the situation is more complex: in addition to direct (mere) negation, e.g., *(ne)závislost* ‘(in)dependence’, *(ne)šťastný* ‘(un)happy’ (Joshi, 2020, p. 86), the prefix *ne-* may lead to a semantic shift (Rechzieglová, 1995, Pavlovič, 2015, Kolářová, 2020; cf. *mysl* ‘sense’ vs. *nesmysl* ‘nonsense’, *ohrožený* ‘threatened’ vs. *neohrožený* ‘fearless’, *patrně* ‘apparently’ vs. *nepatrně* ‘slightly’), referred to here as lexicalized negation. Moreover, N-forms can be used as negative occasional words, for instance in disjunctive syntactic constructions of the type *Zákaz – nezákaz, nalej!* ‘Forbidden or not, pour!’ (Pavlovič, 2015, p. 1365). In order to express meanings slightly different from direct negation, the prefix *bez-* ‘without’ is used in Czech (e.g., *bezdlužnost* ‘state of being without debts’, *bezmocný* ‘powerless’).

The possibility of direct negation in Czech differs according to individual word classes: while all verbs and most adjectives can be negated (Štícha et al., 2013, Pavlovič, 2015, Rusínová & Osolobě, 2016, Šimandl et al., 2016), when it comes to nouns, usually only abstract nouns allow direct negation (Štícha et al., 2013). Furthermore, abstract nouns differ in their use of direct negation depending on their derivational type and semantic category (cf. action, e.g., *(ne)komunikace* ‘non-communication’, and substance, e.g., *komunikace* ‘road’). As for deverbal nouns, only so-called stem nouns (i.e., nouns that in Czech end in *-ní/-tí* and contain a theme suffix of the base verb, e.g., *namítání* – *namítnutí* ‘raising objections’) can be systematically negated if they denote action. In

contrast, so-called root nouns (derived from verbs by various suffixes, including the zero suffix, but not containing a theme suffix, e.g., *námítka* ‘objection’) can be negated only in isolated cases (Dvořák, 2016a, Dvořák, 2016b, Kolářová, 2020). Deadjectival nouns (e.g., *závislost* ‘dependence’) derived from adjectives using N-forms usually allow direct negation, too (Kolářová, 2020; Section 5).

### 3. Negation and Valency

Negative prefixes in the strict sense do not normally affect the argument structure (valency) of the base (Curiel, 2015); however, exceptions exist, both with nouns and adjectives (Haugen, 2013, Kolářová, 2020). Moreover, N-forms can be employed in light verb constructions (Kettnerová et al., 2018) or phrasemes, e.g., *přinést neklid* ‘bring uneasiness’, *plácet nesmysly* ‘talk nonsense’.

When treating negation in a valency lexicon, the following aspects should be considered:

- (i) N-forms that have undergone a semantic shift should undoubtedly be assigned an individual entry (lexeme); the difference in meaning can be accompanied by difference in valency, e.g., *nedůtklivost* ‘touchiness’ has two complements (e.g., *Petrova nedůtklivost vůči kritice* ‘Peter’s touchiness on criticism’) whereas *důtklivost* ‘urgency’ only has one (e.g., *důtklivost jeho výkladu* ‘the urgency of his presentation’);
- (ii) The N-form is the direct negation of the A-form:
  - (ii-a) Typically, A-forms and N-forms have the same valency, and thus separate entries for both A-form and N-form may be regarded as redundant, just as in the case of verbs;
  - (ii-b) Rarely, there is a difference in valency between A-forms and N-forms, cf. Eng. *dependence on* and *independence of / from*, which should then be treated in separate entries;
  - (ii-c) Commonness: N-forms and their valency are usually less common (frequent) than A-forms (e.g., *nevděčný za něco* ‘ungrateful for sth’), however, some N-forms considerably outnumber A-forms (e.g., *nepostradatelný pro někoho* ‘indispensable to sb’).
- (iii) There are cases when only one of the forms (either A-form or N-form) can be a part of a light verb construction or a phraseme, e.g., Cz. *vyhlásit nezávislost* ‘declare independence’ vs. *\*vyhlásit závislost* ‘declare dependence’, Eng. *make sense to / for sb* vs. *make nonsense of sth*.

## 4. Lexicographic Treatment of Negation in NomVallex

### 4.1 Organization of the NomVallex Lexicon

NomVallex is a manually created valency lexicon of Czech nouns and adjectives (Kolářová & Vernerová, 2022), adopting the theoretical framework of Functional Generative Description as its theoretical basis. Its newest version, NomVallex 2.0 (Kolářová et al., 2022), comprises 1,027 lexical units contained in 570 lexemes.

NomVallex follows in the footsteps of the VALLEX lexicon (a valency lexicon of Czech verbs created within the same theoretical framework; Lopatková et al., 2022) and adopts its annotation scheme (see Figure 1). The lexicon entry in both lexicons is called a lexeme; it is an abstract unit associating lexical forms with their lexical units (LUs; Cruse, 1986), i.e., word senses. The entries capture all lexical meanings (senses) of the words included in the lexicon; the lexical units are given their identifier (id), consisting of the string *blu* (i.e., an abbreviation for *basic lexical unit*), letters *v/n/a* differentiating among verbs, nouns, and adjectives, lemma of the word, and the number of the lexical unit (cf. the id *blu-n-závislost-1* of the first lexical unit of the lexeme *závislost* ‘dependence’ in Figure 1). The particular lexical meanings are described by synonyms or roughly paraphrased in a gloss (the *synon* attribute). Aspectual (imperfective vs. perfective) counterparts formed by suffixation, such as *vyzývání<sup>impf</sup>* – *vyzvání<sup>pf</sup>* ‘appealing’ or *ohrožovaný<sup>impf</sup>* – *ohrožený<sup>pf</sup>* ‘threatened’, are treated within a single lexeme. Nouns or adjectives that do not express aspect are assigned the flag *no-aspect*, e.g., *výzva<sup>no-aspect</sup>* ‘appeal’.

The valency properties of a lexical unit are captured in a valency frame, consisting of a set of slots, each standing for one valency complement. Each valency complement is assigned a functor (a label marking the relation of the valency complement to its governing word), and a list of forms determining the surface realization of the complement. The following types of complements may be a part of valency frames: obligatory or optional actants (i.e., ACTor, PATient, ADDRessee, EFFECT, and ORIGIN, e.g., *Petrova<sub>ACT</sub> výzva k pomoci<sub>PAT</sub>* ‘Peter’s appeal for help’, *prodejný mládeži<sub>ADDR</sub>* ‘marketable to young people’), and obligatory free modifications, especially those with the meaning of direction (e.g., *muž povoláný do armády<sub>DIR3</sub>* ‘a man drafted into the army’). In NomVallex, the valency properties of a lexical unit are documented in the example-rich attribute by examples from the Czech National Corpus (two corpora are used, namely SYNv12 (Křen et al., 2023) and Araneum Bohemicum Maximum (Benko, 2015)). The valency frame of the first lexical unit of the noun *závislost* ‘dependence’ and examples documenting it are exemplified in Figure 1.

Extending the basic VALLEX annotation scheme, NomVallex employs new attributes and facilities in order to research phenomena related to part-of-speech specificity of the nouns and adjectives covered (Kolářová et al., 2023). In NomVallex, each lexical unit of an adjective or a noun is assigned a label indicating its derivational type (filled in the attribute type), see Table 1. The label provides information on both part-of-speech membership of the LU (whether it is a noun (N) or an adjective (A)) and its derivational base (whether it is deverbal (DV), deadjectival (DA), denominal (DN) or

primary (P)); for more details on the labels, see Kolářová et al., 2023, e.g., the noun *závislost* ‘dependence’ is given the label N-DA-3-lost.<sup>1</sup> The structure of the NomVallex 2.0 lexicon, reflecting the basic derivational categories captured in the lexicon, is presented in Table 1.

**Table 1:** The structure of the NomVallex 2.0 lexicon

Part-of-speech category	Derivational category	Derivational type	Lexical units	Lexemes
<b>Nouns</b>	deverbal	N-DV-...	505	248
	deadjectival	N-DA-...	247	158
<b>Adjectives</b>	deverbal	A-DV-...	179	116
	denominal	A-DN-...	28	14
	deadjectival	A-DA-...	6	6
	primary	A-P	62	28
<b>Total</b>			1,027	570

To enable analysis of the relationship between the valency behavior of base words and their derivatives, lexical units of nouns and adjectives in NomVallex are linked to their respective base lexical units (contained either in NomVallex itself or, in the case of verbs, in the VALLEX lexicon). Derivationally related lexical units of nouns and adjectives are linked to each other (or to their base verbs in VALLEX) by means of two attributes which keep both directions, namely (i) the attribute *derivedFrom* (it provides a link from a particular LU to its base LU), and (ii) the attribute *derivedLUs* (being dynamically generated according to the latest data, it captures links to all LUs derived from the base LU).

The treatment of negative forms of nouns and adjectives presented in this paper is a new development in the lexicon data that will feature in the future published version (in 2024).

## 4.2 Negation-Related Attributes in NomVallex

As a manually annotated lexicographic resource, the NomVallex lexicon aspires to treat negation of nouns and adjectives as accurately as possible. All lexical units of nouns and adjectives contained in NomVallex are examined with respect to whether or not they can be used in N-forms expressing direct negation of the A-forms. The N-forms are searched for in the two corpora used for documenting the A-forms, namely in SYNv12 and Araneum Bohemicum Maximum. An annotator then looks through the occurrences found and manually annotates them in the data of the lexicon,

<sup>1</sup> In particular, the label N-DA-3-lost indicates a noun derived from an adjective (i.e., *závislost* ‘dependence’ < *závislý* ‘dependent’) that is derived from an active participle (this verbal form being marked here by number 3; i.e., *závislý* ‘dependent’ < *závisel* ‘depended’). Moreover, as the mnemonic aid, the label contains the string in which this derivational type of noun ends, namely *-lost*.

differentiating between particular lexical units (senses) of the N-forms, if necessary. If a word among the N-forms turns out to have its own (lexicalized) meaning, then it is assigned an individual entry (lexeme; e.g., *nezvyklý* ‘unusual’).<sup>2</sup>

In order to tackle the issues specified in Section 3, N-forms in NomVallex that directly negate lexical units expressed in A-forms and that do not differ in valency are captured within the entries for A-forms. In extension of the basic annotation scheme described in Section 4.1, the negation-related attributes specified in Table 2 are employed, if relevant (the treatment of the N-form *nezávislost* ‘independence’ is exemplified in Figure 1).

```
* ZÁVISLOST [NomVallex.txt]
-lexeme-affirm-freq: 187612
-lexeme-affirm-i-p-m: 30,08
-lexeme-neg-freq: 98257
-lexeme-neg-i-p-m: 15,75
~ no-aspect: závislost [blu-n-zavislost-1]
+ ACT2, pos PATna+6
-synon: odkázanost
-derivedFrom: blu-a-zavisly-1
-example-rich: ACT+PAT: Zdejší zásoby závislost Slovenska.ACT na dovozu.PAT plynu z Ruska nezmění. • Za své selhání považovali přílišnou závislost dětí.ACT na matce.PAT. • S tím, jak se zvyšuje počet vysokoškolských studentů, se prodlužuje období jejich.ACT závislosti na rodičích.PAT. • Vojenských alternativ měli poskrovnu, jejich.ACT závislost na Západu.PAT sílíla
-type: N-DA-3-lost
-neg-lemmas: no-aspect: nezávislost
-neg-id: blu-n-zavislost-1:neg
-neg-commonness: standard
-neg-type: direct negation
-neg-example: ACT+PAT: To by umožnilo její.ACT nezávislost na politických stranách.PAT. • Předložený studijní grant, který by studentovi měsíčně přispíval na studia a zvyšoval tak jeho.ACT nezávislost na rodičích.PAT, se pro změnu neopírá o žádnou finanční rozvahu. • Zvýší se také jeho.ACT nezávislost na sousedním Rusku.PAT. • Půjčky se tradičně užívají v severských zemích a v Nizozemsku: pro tyto země je totiž charakteristická brzká nezávislost dětí.ACT na rodičích.PAT. • V USA těžba plynu a ropy z břidelic znamená faktickou nezávislost Spojených států.ACT na importech.PAT.
-neg-lvc: blu-v-vyhlasit-vyhlasovat-4
```

Fig. 1: The N-form *nezávislost* ‘independence’ is captured within the entry for the A-form *závislost* ‘dependence’

First, the attribute *neg-type* specifies whether the N-form represents direct negation or lexicalized negation. The attribute *neg-type* also captures cases of lemmas that do not express word-level negation, differentiating between so-called *negativum tantum* (the term used in Czech terminology for words that start with the string *ne-* but that in present-day Czech have no meaning without the string, e.g., *nedočkavý* ‘impatient’, *nenávist* ‘hatred’) and words considered not to be able to occur in the N-forms for various reasons (the value *inapplicable* is used in this case).

<sup>2</sup> As NomVallex specializes in non-verbal predicates endowed with valency, the lexicographic treatment of negative forms of Czech nouns and adjectives covers only a small part of Czech lexis. As for nouns, in the SYNv12 corpus there are 8,876 nouns, the lemmas of which start with the string *ne-* (including words that have no relation to negation, such as *nebe* ‘sky’ or *nerv* ‘nerve’), whereas NomVallex only features slightly more than 400 noun lemmas. As for adjectives, the SYNv12 corpus contains 20,207 lemmas of adjectives that have at the 11<sup>th</sup> position of their morphological tag the label N used for negation, whereas NomVallex captures fewer than 200 adjectival lemmas.

N-forms representing direct negation are assigned the particular negative lemma(s) in the ‘neg-lemmas’ attribute. With lexicalized negation, negativum tantum or inapplicable N-form(s), the value of the ‘neg-lemmas’ attribute is ‘NA’. If a lexical unit has no hit for its N-form in the corpora but it could theoretically be used, the value of its ‘neg-lemmas’ attribute is the appropriate N-form, and the ‘neg-commonness’ attribute is assigned the value ‘NO example’. Nevertheless, it is sometimes difficult to distinguish clearly between the two cases of no occurrence of the particular N-form.

**Table 2:** Negation-related attributes of lexical units in the NomVallex lexicon

Attribute	Value	Example
<b>neg-type</b>	direct negation	<i>nezávislost</i> ‘independence’
	lexicalized negation	<i>neohrožený</i> ‘fearless’
	negativum tantum	<i>nenávist</i> ‘hatred’
	inapplicable	<i>nezkouška</i> ‘non-exam’
<b>neg-lemmas</b>	lemma(s) of the N-form	<i>nezávislost</i> ‘independence’
	NA	* <i>nedomněnka</i> ‘unsurmise’
<b>neg-commonness</b>	NO example	<i>neapelování</i> ‘non-appealing’
	rare	<i>nezvědavost</i> ‘incuriosity’
	rare, marked	<i>nepodmínka</i> ‘non-condition’
	standard	<i>nezávislost</i> ‘independence’
	more frequent than the A-form	<i>nepostradatelný</i> ‘indispensable’
<b>neg-example</b>	annotated corpus examples	<i>To by umožnilo její<sub>ACT</sub> nezávislost na politických stranách<sub>PAT</sub>.</i> ‘It would enable her independence from political parties.’
<b>neg-id</b>	identifier of the N-form	blu-n-závislost-1:neg
<b>neg-lvc</b>	identifier(s) of light verb(s) with which the N-form forms a predicate	blu-v-vyhlásit-vyhlášovat-4 (cf., <i>vyhlásit</i> ‘declare’ + the N-form <i>nezávislost</i> ‘independence’)

The other values of the ‘neg-commonness’ attribute express the subjective annotator’s evaluation of how common the N-forms are with respect to the corresponding A-forms (primarily examples with valency are taken into account); the values include ‘rare’, ‘rare, marked’, ‘standard’, and ‘more frequent than the A-form’ (see Section 5 for more details). The values of the ‘neg-commonness’ attribute are somewhat vague (i.e., not numerical), but thanks to manual annotation they are able to capture the differences between particular senses of a lexeme.

In addition to the ‘neg-commonness’ attribute, the NomVallex data are also automatically supplemented with attributes providing the absolute and relative corpus frequency of both

affirmative and negative forms of the words included (the SYNv12 corpus is used). However, the corpus frequency (by definition) cannot distinguish between individual lexical units (senses) (e.g., *prodejnost výrobku* ‘marketability of the product’ vs. *prodejnost politika* ‘venality of a politician’), homonyms (e.g., *výklad* ‘explanation’ vs. *výklad* ‘shop window’) or reflexive and non-reflexive lexemes (cf., *zážitek nedorozumění se s někým* ‘experience of not understanding each other’ and *malé nedorozumění* ‘slight misunderstanding’). Therefore, despite being ‘objective’, the information value of corpus frequency is in fact in many cases only approximate. As the frequency data are not relevant to individual lexical units, attributes covering corpus frequency are listed at the level of lexemes, not lexical units (attributes capturing corpus frequency of the affirmative and negative lemmas of the lexeme *závislost* ‘dependency’, i.e., ‘lexeme-affirm-freq’, ‘lexeme-neg-freq’, ‘lexeme-affirm-i-p-m’, and ‘lexeme-neg-i-p-m’, are exemplified in Figure 1).<sup>3</sup>

The attribute ‘neg-example’ contains manually annotated examples, handpicked from the corpora used; the examples should primarily document the valency properties of the particular N-forms.

In order to tackle the cases when only one of the forms (either the A-form or the N-form) can be a part of a phraseme or a light verb construction (lvc; e.g., *vyhlásit nezávislost* ‘declare independence’), in addition to the A-form’s identifier (e.g., *blu-n-závislost-1*), also each N-form annotated within a lexical unit together with the A-form is given its own identifier (the ‘neg-id’ attribute, e.g., *blu-n-závislost-1:neg*). As a result, the particular N-form and the light verb can be interlinked (e.g., *nezávislost* ‘independence’ and the light verb *vyhlásit* ‘declare’). If relevant, the identifier(s) of the light verb(s) with which the N-form forms a predicate are specified in the ‘neg-lvc’ attribute.

## 5. A Case Study: Negative Forms of Deadjectival Nouns in NomVallex and Their Valency

The manually annotated negation-related attributes (Section 4.2) provide an overall picture of the use of negative forms of nouns and adjectives contained in the NomVallex lexicon. To illustrate the use and distribution of the negative forms, including their valency properties, we focus here on one derivational category annotated in accordance with the guidelines described in Section 4, namely on Czech deadjectival nouns (in total 247 lexical units in 158 lexemes, see Table 1; e.g., *závislost* ‘dependence’ < *závislý* ‘dependent’).

We assume that the values of the ‘neg-commonness’ attribute in particular (described below) can serve as a good indicator of how frequent or common the N-forms are compared to their corresponding A-forms. The distribution of the values assigned in our data to deadjectival nouns is presented in Table 3.

<sup>3</sup> Nevertheless, homonyms and reflexive vs. non-reflexive words captured in the lexicon are not provided with the corpus frequency attributes because they would not be relevant in these cases.

**Table 3:** The ‘neg-commonness’ attribute of lexical units of deadjectival nouns in NomVallex

N-form: direct negation	Value of the attribute	Number of lexical units	
		absolute	relative
No	not present	34	13,8
Yes	standard	82	33,2
	rare	75	30,4
	rare, marked	6	2,4
	more frequent than the A-form	3	1,2
	NO example	47	19,0
<b>Total</b>		247	100%

While nouns in the first category (labeled with the value ‘not present’) clearly differ from typical A-form – N-form pairs and are mentioned here for the sake of completeness only, all the other categories represent cases of direct negation of Czech deadjectival nouns:

- ‘not present’: The lexical units in question are represented by (i) lexicalized N-forms (e.g., *neohroženost* ‘fearlessness’), (ii) nouns that in present-day Czech have no meaning without the string *ne-* (e.g., *nedočkavost* ‘impatience’), (iii) nouns to which the N-form is not relevant for some other reason, for example because their A-form is lexicalized or is employed in a phraseme and thus cannot be negated (e.g., *drahota* used in the phraseme *dělat drahoty* ‘make a fuss’). Their attribute ‘neg-lemmas’ receives the value ‘NA’ (see Section 4.2), and thus the ‘neg-commonness’ attribute is not relevant to them.
- ‘standard’: Being quite frequent, these N-forms use the same valency as their A-form counterparts, such as the N-form *nezávislost* ‘independence’ and the A-form *závislost* ‘dependence’, both amply documented with Actor expressed in the genitive or possessive forms and with Patient in a prepositional phrase introduced by *na* ‘on / from’ (see Figure 1).
- ‘rare’: These N-forms are usually documented in few examples, being considerably less frequent than their corresponding A-forms. The N-forms are only rarely modified by valency complements (e.g., *nehrdost* ‘non-pride’); moreover, usually also some of the complements (or some of their forms) are typically not documented at all, although they probably could be used.
- ‘rare, marked’: These N-forms are not only rare, but the examples found in the corpora used even seem to be ‘marked’ or ‘odd’ (e.g., *nevlastnost* ‘non-property’).
- ‘more frequent than the A-form’: In isolated cases, N-forms far outnumber the corresponding A-forms (e.g., *nedotknutelnost* ‘untouchability’).
- ‘NO example’: These N-forms are not documented in any example in the corpus data. However, they could theoretically be used (e.g., *neuraženost* ‘non-pique’).

The distribution of the values of the ‘neg-commonness’ attribute in our data suggests that the use of N-forms of Czech deadjectival nouns more or less corresponds to tendencies described in previous studies (see e.g., Kolářová, 2020). Most deadjectival nouns seem to use their N-forms for expressing direct negation. However, only about one third of the nouns in our data use their N-forms fairly frequently and with the same valency as the corresponding A-forms (the ‘standard’ N-forms in Table 3). Leaving aside cases of non-expression of valency complements, overall, only slight differences in valency between A-forms and N-forms are attested in our data. The prepositional phrases of complements other than Actor are usually kept with the N-forms, and they are sometimes preferred to prepositionless cases; for instance, while the Patient of the A-form of the noun *škodlivost* ‘harmfulness’ is documented in a prepositionless dative (e.g., *jeho škodlivost lidskému zdraví* ‘its harmfulness to human health’), it is not attested with the N-form *neškodlivost* ‘non-harmfulness’. Instead, both A-form and N-form (*ne*)*škodlivost* ‘(non-)harmfulness’ can be modified by the Patient in the prepositional phrase introduced by *pro* ‘for’ (e.g., (*ne*)*škodlivost pro lidské zdraví* ‘(non-)harmfulness to human health’). Exceptionally, a complement of the N-form takes on a form that is not used with the corresponding A-form at all; for example, while both the A-form *libost* ‘pleasure’ and the N-form *nelibost* ‘displeasure’ can be modified by the Patient in the prepositional phrase introduced by *nad* ‘at’ (e.g., (*ne*)*libost nad jejich počínáním* ‘(dis)pleasure at their actions’), only the N-form *nelibost* ‘displeasure’ can be modified by the Patient in the prepositional phrase introduced by *s* ‘with’, see (1).

(1) neskrývala	nelibost	s	výroky
not-hide- PST-1SG	displeasure-ACC	with	statements-INS
nového	vedení		
new-GEN	management-GEN		
‘(she) did not hide displeasure with the statements of the new management’			

Cases in which N-forms of Czech deadjectival nouns are employed in light verb constructions are relatively rare in our data, e.g., *vyhlásit nezávislost* ‘declare independence’, *vyjádřit nespokojenost* ‘express discontent’, *projevit nelibost* ‘show displeasure’, *vyvolat nevolnost* ‘cause sickness’. N-forms of deadjectival nouns can also be used in idioms such as *pracovní neschopnost* ‘incapacity for work’.

## 6. Conclusion

Various aspects of word-level negation in Czech nouns and adjectives pose a challenge to lexicographers. In this paper, we have specified some of the issues to be resolved when treating word-level negation in a valency lexicon. Working on the assumption that direct negation of Czech nouns and adjectives should be treated within entries for the corresponding affirmative forms, we have sketched out annotation guidelines and the negation-related attributes applied to the data of the NomVallex lexicon. We have presented preliminary results of the annotation, illustrating these through the case of deadjectival nouns. We believe that this lexicographic treatment will facilitate further research into negation in Czech nouns and adjectives.

## References

- Benko, V. (2015). *Araneum Bohemicum Maximum, verze 15.04*. Ústav Českého národního korpusu FF UK, Praha. Available at <https://www.korpus.cz>.
- Cruse, D. A. (1986). *Lexical Semantics*. Cambridge University Press.
- Curiel, M. M. (2015). Negation. In P. O. Müller et al. (Eds.), *Word-Formation. An International Handbook of the Languages of Europe. Volume 2* (pp. 1351–1359). De Gruyter Mouton.
- Dvořák, V. (2016a). Dějové substantivum. In P. Karlík et al. (Eds.), *Nový encyklopedický slovník češtiny* (pp. 309–315). Nakladatelství Lidové noviny.
- Dvořák, V. (2016b). Verbální substantivum. In P. Karlík et al. (Eds.), *Nový encyklopedický slovník češtiny* (pp. 1948–1953). Nakladatelství Lidové noviny.
- Hajič, J. et al. (2003). PDT-VALLEX: Creating a Large-coverage Valency Lexicon for Treebank Annotation. In *Proceedings of The Second Workshop on Treebanks and Linguistic Theories* (pp. 57–68). Vaxjo University Press.
- Haugen, T. A. (2013). Adjectival valency as valency constructions. Evidence from Norwegian. *Constructions and Frames*, 5(1), 35–68. <https://doi.org/10.1075/cf.5.1.02hau>
- Herbst, T. et al. (2004). *A valency dictionary of English: a corpus-based analysis of the complementation patterns of English verbs, nouns, and adjectives*. Walter de Gruyter.
- Joshi, S. (2020). Affixal negation. In V. Déprez & M. T. Espinal (Eds.), *The Oxford Handbook of Negation* (pp. 75–88). Oxford University Press.
- Kettnerová, V., Lopatková, M., Bejček, E., & Barančíková, P. (2018). Enriching VALLEX with Light Verbs: From Theory to Data and Back Again. *The Prague Bulletin of Mathematical Linguistics*, 111, 29–56. <https://doi.org/10.2478/pralin-2018-0006>
- Kolářová, V. (2020). Vztah afirmativní a negované formy adjektiv a substantiv z hlediska jejich valence. *Prace Filologické*, 75(1), 293–312. <https://doi.org/10.32798/pf.662>
- Kolářová, V., Kettnerová, V., & Mírovský, J. (2023). Through Derivational Relations to Valency of Non-verbal Predicates in the NomVallex Lexicon. *Jazykovedný časopis*, 74(1), 182–192. <https://doi.org/10.2478/jazcas-2023-0036>
- Kolářová, V., & Vernerová, A. (2022). NomVallex: A Valency Lexicon of Czech Nouns and Adjectives. In *Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)* (pp. 1344–1352). European Language Resources Association.
- Kolářová, V., Vernerová, A., & 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. Available at <http://hdl.handle.net/11234/1-4663> and at <https://ufal.mff.cuni.cz/nomvallex>.
- Kováříková, D. (2011). Gramatická kategorie negace. In F. Čermák (Ed.), *Korpusová lingvistika Praha 2011. 2: Výzkum a výstavba korpusů* (pp. 273–280). Nakladatelství Lidové noviny.
- Kováříková, D., Chlumská, L., & Cvrček, V. (2012). What Belongs in a Dictionary? The Example of Negation in Czech. In R. V. Fjeld & J. M. Torjusen (Eds.), *Proceedings of the 15th EURALEX International Congress* (pp. 822–827). Department of Linguistics and Scandinavian Studies, University of Oslo.
- Křen, M. et al. (2023). *Korpus SYN, verze 12 ze 29. 12. 2023*. Ústav Českého národního korpusu FF UK, Praha. Available at <https://www.korpus.cz>.
- Lopatková, M. et al. (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. Available at <http://hdl.handle.net/11234/1-4756>.
- Lotko, E. (1973). *Lexikální negace v současné češtině*. Státní pedagogické nakladatelství.
- Pavlovič, J. (2015). Negation in the Slavic and Germanic languages. In P. O. Müller et al. (Eds.), *Word-Formation. An International Handbook of the Languages of Europe. Volume 2* (pp. 1360–1373). De Gruyter Mouton.

Rechzieglová, A. (1995). *On Negation in Czech*. Euroslavica.

Rusínová, Z., & Osolobě, K. (2016). Prefixace. In P. Karlík et al. (Eds.), *Nový encyklopedický slovník češtiny* (pp. 1394–1395). Nakladatelství Lidové noviny.

Svozilová, N., Prouzová, H., & Jirsová, A. (2005). *Slovník slovesných, substantivních a adjektivních vazeb a spojení*. Academia.

van Son, C., van Miltenburg, E., & Morante, R. (2016). Building a Dictionary of Affixal Negations. In *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics (ExProM)* (pp. 49–56). Osaka, Japan.

Šimandl, J. et al. (2016). *Slovník afixů užívaných v češtině*. Karolinum.

Štícha, F. et al. (2013). *Akademická gramatika spisovné češtiny*. Academia.

Zimmer, K. E. (1964). Affixal negation in English and other languages: An investigation of restricted productivity. *Word*, 20, Issue sup. 1, Monograph No. 5.

## Acknowledgements

This work was supported by the Czech Science Foundation (project No. 22-20927S). The work described herein has been using data and tools provided by the LINDAT/CLARIAH-CZ Research Infrastructure (<https://lindat.cz>), supported by the Ministry of Education, Youth and Sports of the Czech Republic (project No. LM2023062).

## Contact information

### Veronika Kolářová

Institute of Formal and Applied Linguistics, Faculty of Mathematics and Physics, Charles University, Czech Republic  
kolarova@ufal.mff.cuni.cz

### Jiří Mírovský

Institute of Formal and Applied Linguistics, Faculty of Mathematics and Physics, Charles University, Czech Republic  
mirovsky@ufal.mff.cuni.cz