PBML

The Prague Bulletin of Mathematical Linguistics
NUMBER 111 OCTOBER 2018 29–56

Enriching VALLEX with Light Verbs: From Theory to Data and Back Again

Václava Kettnerová, Markéta Lopatková, Eduard Bejček, Petra Barančíková

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

Abstract

This paper summarizes results of a theoretical analysis of syntactic behavior of Czech light verb constructions and their verification in the linguistic annotation of a large amount of these constructions. The concept of LVCs is based on the observation that nouns denoting actions, states, or properties have a strong tendency to select semantically underspecified verbs, which leads to a specific rearrangement of valency complementations of both nouns and verbs in the syntactic structure. On the basis of the description of deep and surface syntactic properties of LVCs, a formal model of their lexicographic representation is proposed here. In addition, the resulting data annotation, capturing almost 1,500 LVCs, is described in detail. This annotation has been integrated in a new version of the VALLEX lexicon, release 3.5.

1. Introduction

Light verb constructions (LVCs) pose a serious challenge for both theoretical linguistics and NLP tasks due to their syntactic complexity. The major challenges raised by LVCs can be overcome by a lexicographic representation allowing for their efficient handling in both theoretical and computational linguistics. Developing a formal model of such representation thus represents a crucial task of the current lexicography.

In this paper, we present a formal model of the lexicographic description of LVCs designed for the valency lexicon of Czech verbs VALLEX, summarizing findings partially presented esp. in Kettnerová and Lopatková (2015); Kettnerová et al. (2016); Kettnerová (2017); Kettnerová and Lopatková (2017b), and in Kettnerová and Lopatková

© 2018 PBML. Distributed under CC BY-NC-ND. Corresponding author: kettnerova@ufal.mff.cuni.cz
Cite as: Václava Kettnerová, Markéta Lopatková, Eduard Bejček, Petra Barančíková. Enriching VALLEX with Light Verbs: From Theory to Data and Back Again. The Prague Bulletin of Mathematical Linguistics No. 111, 2018, pp. 29–56. doi: 10.2478/pralin-2018-0006.

(2017a). This model, based on a thorough theoretical research into Czech LVCs and grounded in an in-depth analysis of corpus data using the Functional Generative Description (FGD, see esp. Sgall et al., 1986), has been applied in an extensive annotation of Czech LVCs allowing us to verify theoretical adequacy of the adopted postulates and their further modification. This annotation has been integrated in the VALLEX lexicon, release 3.5.

LVCs represent a type of complex predicates where two syntactic elements – a light verb and a predicative noun, adjective, adverb (or verb esp. in Asian languages) – function together as a single predicative unit. The syntactic structure of an LVC is then determined by both the light verb and the other predicative element. For example, the structure with the light verb *dát* 'to give' is determined not solely by the verb, which provides three valency complementations, namely ACT (*matka* 'mother'), ADDR (*děti* 'children'), and CPHR (*příkaz* 'order'), see example (1), but also by the noun *příkaz* 'order', which contributes PAT (*uklidit* 'to clean up') to the structure.[1]

(1) *Matka*$_{\text{ACT}}$ *dala dětem*$_{\text{ADDR}}$ *příkaz*$_{\text{CPHR}}$ *uklidit*$_{\text{PAT}}$ *pokoj.*
'Mother gave children an order to clean their room up.'

The question how valency complementations of light verbs and predicative elements participate in the syntactic structure formation of LVCs is a central issue of any syntactic theory attempting to provide their comprehensive analysis. Despite being addressed in many theoretical frameworks, see e.g., argument merger formulated within the Government Binding theory (Grimshaw and Mester, 1988), argument fusion (Butt, 2010) and argument composition within the Lexical-Functional Grammar (Hinrichs et al., 1998), and the study by Alonso Ramos carried out within the Meaning $\leftrightarrow$ Text Theory (Alonso Ramos, 2007), this issue is still far from being clear. The main difficulty in arriving at a more uniform analysis of LVCs lies in the fact that a definitional characterization allowing for their cross-linguistic identification is still missing. In some analyses, LVCs are then mixed up with control constructions, or even with auxiliary verbs – such misinterpretations clearly lead to inconsistent conclusions (see esp. Butt and Geuder, 2001).

In this paper, we aim at contributing to better understanding the syntactic structure formation of Czech LVCs, making use of the theoretical framework of FGD. We limit our focus to LVCs composed of light verbs and predicative nouns expressed as their direct object, as these LVCs are the most central and frequent ones in Czech. We demonstrate that the syntactic formation of these LVCs is compositional, namely that the syntactic structure of an LVC can be inferred from syntactic properties of the predicative noun and the light verb forming the given LVC on a rule basis. Further, we

[1] Compare with the syntactic structure with the full verb *dát* consisting of ACT (*matka* 'mother'), ADDR (*děti* 'children'), and PAT (*čokoláda* 'chocolate'), see the following example:
Matka$_{\text{ACT}}$ *dala dětem*$_{\text{ADDR}}$ *čokoládu*$_{\text{PAT}}$.
'Mother gave children chocolate.'

show that the coreference relation between valency complementations of predicative nouns and light verbs represents a key characteristic of LVCs.

As has been already reflected in the literature (see e.g., Radimský, 2010), an intuitive clue lying in paraphrasability of LVCs by single verbs does not represent a reliable criterion for their identification as there is a number of broadly accepted LVCs without a suitable paraphrase (e.g., *udělat dojem* 'to make an impression', *nést vinu* 'to be at fault', *mít názor* 'to have an opinion'); on the contrary, some full verb collocations can be paraphrased by single verbs (e.g., *uložit do hrobu* 'to lay to rest; lit. to lay into the grave' – *pohřbít* 'to tomb'). Therefore we consider the syntactic compositionality together with the coreference relation between verbal and nominal valency complementations as a definitional criterion for delimiting LVCs in our research.

The paper is structured as follows. First, the valency lexicon of Czech verbs, VALLEX, and its theoretical background – the valency theory of FGD – are introduced (Section 2). Second, the results of the theoretical analysis of Czech LVCs are summarized (Section 3). On the basis of these results, a formal model of the lexicographic representation of LVCs is proposed (Section 4). Third, the application of this model to the annotation of a large amount of linguistic data is described and the resulting annotated data are characterized (Section 5). The last section provides a short summary.

2. VALLEX and the Valency Theory of FGD

VALLEX, a valency lexicon of Czech verbs, takes the Functional Generative Description (FGD) as its theoretical background. FGD represents a dependency oriented framework which adopts a stratificational approach to the language description (Sgall et al., 1986). One of the main concept of FGD is represented by the tectogrammatical layer – the deep syntactic layer. The core of this layer is represented by valency, an ability of a word to open a certain number of valency positions for other dependent units. The valency theory of FGD has been elaborated from the 70th esp. by Jarmila Panevová (see esp. Panevová, 1974–75, 1980, 1994) and applied in several valency lexicons: *VALLEX*, a valency lexicon of Czech verbs (Lopatková et al., 2016),[2] *PDT-Vallex*, a valency lexicon linked to the family of Prague Dependency Treebanks (Urešová, 2011),[3] and *EngVallex*, a valency lexicon representing the part of the annotation of the Prague Czech-English Dependency Treebank.[4]

In the valency theory of FGD, actants and free modifications are distinguished. Actants (be they obligatory, or optional) represent valency complementations char-

[2] http://ufal.mff.cuni.cz/vallex (http://hdl.handle.net/11234/1-2307)

[3] http://ufal.mff.cuni.cz/pdt-vallex-valency-lexicon-linked-czech-corpora (http://hdl.handle.net/11858/00-097C-0000-0023-4338-F)

[4] http://ufal.mff.cuni.cz/engvallex-english-valency-lexicon-linked-corpora (http://hdl.handle.net/11858/00-097C-0000-0023-4337-2)

acterizing the word in a unique way; as such they have to be listed in its valency frame. Five actants are recognized: ACT, PAT, ADDR, ORIG, and EFF. A higher number of free modifications is determined on the basis of their semantic features (e.g., temporal, spatial, causal, etc.); only obligatory ones characterize a word in a unique way and thus they have to be part of its valency frame as well (see esp. Panevová, 1994; Lopatková and Panevová, 2006). Morphemic forms of actants are determined by a governing word (as such they have to be indicated in its valency frame) while forms of free modifications stem from their semantic type (as a result, they do not have to be provided in the valency frame) (Lopatková and Panevová, 2006).

The VALLEX lexicon, attempting to provide a comprehensive description of valency behavior of Czech verbs, represents the most elaborated lexicon developed within FGD. For purposes of the description of language phenomena at the lexicon-grammar interface, this lexicon has been divided into two components: a lexical part (the data component), providing information specific to individual lexical units, and a grammar part (the grammatical component), capturing regular patterns of Czech verbs in the form of formal rules (thus being a part of the overall grammar of Czech).

Data component provides the information on valency structure of Czech verbs in their individual senses. The key organizing concept of this lexicon is represented by the *lexeme*, an abstract twofold unit associating lexical forms of a verb with its lexical units (individual senses). Each *lexical unit* is assigned the syntactic and semantic information. VALLEX stores more than 6,760 lexical units of verbs (counting aspectual counterparts separately 10,900 lexical units) contained in more than 2,730 lexemes. These lexemes are represented by almost 4,600 verb lemmas.

The crucial information on the valency structure of individual lexical units is provided in the form of valency frame. Valency frame is modeled as a sequence of valency slots, each slot standing for one valency complementation. Each slot of a valency complementation is characterized by a functor (a syntactico-semantic label marking the relation of the valency complementation to its governing verb, as e.g., ACT, ADDR, PAT, LOC etc.) and by the information on obligatoriness. In addition, the information on morphemic forms is provided for actants, indicating possible surface syntactic expression of the given complementation in active, unreciprocal and irreflexive constructions.

Each lexical unit can be described by other relevant syntactic and syntactico-semantic information, e.g., on control, reflexivity, reciprocity, diatheses, alternations, syntactico-semantic class membership.

Grammar component stores grammatical rules which instantiated by the information provided by the data component of the lexicon allow for obtaining all possible surface syntactic manifestations of lexical units of verbs, namely their passive, reciprocal and reflexive structures.

As we show below, the division into the lexical and grammar component has appeared to be relevant also for the description of LVCs as they represent a typical language phenomenon bridging these two parts of the language description.

3. Theoretical Analysis of LVCs

In this section, we provide a comprehensive theoretical account of both deep (Section 3.1) and surface structure formation of LVCs (Section 3.2), grounded in FGD.

3.1. Deep Syntactic structure of LVCs

The deep syntactic structure of LVCs consists of valency complementations of both predicative nouns (Section 3.1.1) and light verbs (3.1.2). In addition, it is characterized by coreference between valency complementations of light verbs and predicative nouns (Section 3.1.4).

3.1.1. Valency Frames of Predicative Nouns

As predicative nouns, deverbal nouns (e.g., *plán* 'plan', *hněv* 'rage', *nedůvěra* 'distrust', *nenávist* 'hate', *svědectví* 'evidence'), deadjective nouns (e.g., *vděčnost* 'gratitude', *žárlivost* 'jealousy'), and primary nouns (e.g., *láska* 'love') occur. Predicative nouns denote actions, states (deverbal and primary nouns), and properties (deadjective nouns). Deverbal and deadjective predicative nouns typically inherit the valency structure from their base verbs and adjectives, respectively. The valency structure of primary nouns can be accounted for on the basis of the valency structure of verbs with corresponding meanings (e.g., *láska* 'love' and *milovat* 'to love'). The number and type of valency complementations of predicative nouns (i.e., their functors and obligatoriness) thus remain the same (see esp. Kettnerová et al., 2017). Only morphemic forms of their complementations usually undergo changes, reflecting their usage in nominal structures. Compare, e.g., the valency frame of the predicative noun *obava* 'fear' (5) and its nominal structure (7) with the frame of its base verb *obávat se* 'to fear; to be afraid' (2) and its verbal structure (4).

(2) *obávat se* 'to fear; to be afraid': $\mathsf{ACT}_{\mathsf{nom}}\ \mathsf{PAT}_{\mathsf{gen,inf,dcc}}$

(3)
$$\begin{array}{lcl} \mathsf{ACT}_{\mathsf{V}} & \Leftrightarrow & \text{'Experiencer'} \\ \mathsf{PAT}_{\mathsf{V}} & \Leftrightarrow & \text{'Stimulus'} \end{array}$$

(4) *Věřící lidé*$_{\mathsf{ACT.nom}}$ *se méně obávají smrti*$_{\mathsf{PAT.gen}}$.
'Believers$_{\mathsf{ACT}}$ are less afraid of death$_{\mathsf{PAT}}$.'

(5) *obava* 'fear': $\mathsf{ACT}_{\mathsf{gen,pos}}\ \mathsf{PAT}_{\mathsf{před+instr,z+gen,inf,dcc}}$

(6)
$$\begin{array}{lcl} \mathsf{ACT}_{\mathsf{N}} & \Leftrightarrow & \text{'Experiencer'} \\ \mathsf{PAT}_{\mathsf{N}} & \Leftrightarrow & \text{'Stimulus'} \end{array}$$

(7) *obavy věřících lidí*$_{\text{ACT.gen}}$ *ze smrti*$_{\text{PAT.z+gen}}$
'believers'$_{\text{ACT}}$ fear of death$_{\text{PAT}}$'

Each valency complementation of a predicative noun corresponds to a semantic participant of the situation denoted by the noun. Semantic participants of a noun are typically identical to the participants characterizing its base verb (or a semantically corresponding verb). For example, the valency complementations ACT and PAT of the predicative noun *obava* 'fear' are mapped onto the semantic participant 'Experiencer', a sentient entity experiencing fear, and 'Stimulus', evoking the given emotion, respectively, as complementations of its base verb *obávat se* 'to fear; to be afraid', see (6) and (3), respectively.[5]

3.1.2. Valency Frames of Light Verbs

Light verbs, as semantically impoverished verbs (e.g., Jespersen, 1965; Grimshaw and Mester, 1988; Butt and Geuder, 2001), denote only general semantic scenarios. They typically inherit valency characteristics from respective full verb counterparts – while one of their valency position is reserved for predicative nouns (this position is labeled with the CPHR functor in FGD), other valency complementations acquire semantic specifications in LVCs via coreference with valency complementations of predicative nouns (see esp. Alonso Ramos, 2007); Section 3.1.4 discusses this feature in more detail.[6] The only exception represented by 'Causator' is discussed below in this Section.

Let us compare the valency frame of the full verb *mít* 'to have' (8) and the frame of the light verb *mít* 'to have' (10) and their example sentences (9) and (12), respectively. The full verb refers to a possession of an object by an owner (its valency complementations are mapped onto the semantic participants: ACT onto 'Owner' and PAT onto 'Possession'). In contrast, the light verb, denoting a general scenario, does not have any semantic participants semantically specifying its valency complementations. Their complementations are semantically specified just in LVCs, by entering in coreference with nominal complementations. For example, in the LVC *mít obavu* 'to be afraid; lit. to have a fear', the verbal ACT attains its semantic saturation via coreference with the nominal ACT, namely it refers to 'Experiencer', see the valency frame (5) and the mappings (6) and (11).

(8) *mít* 'to have': ACT$_{\text{nom}}$ PAT$_{\text{acc}}$

(9) *Petr*$_{\text{ACT.nom}}$ *má nový dům*$_{\text{PAT.acc}}$.
'Peter$_{\text{ACT}}$ has a new house$_{\text{PAT}}$.'

[5] In examples below, the ⇔ arrows indicate the mapping between valency complementations and semantic participants; the ≡ sign indicates predicative nouns filling the CPHR position; the ↔ arrows are reserved for coreferential relations between verbal and nominal valency complementations.

[6] The possible reduction of valency frames of light verbs is discussed in Kettnerová and Lopatková (2013).

(10) *mít* 'to have': ACT_{nom} CPHR_{acc}

(11) *mít obavu* 'to be afraid':

		ACT_{V}	$\leftrightarrow$	ACT_{N}	$\Leftrightarrow$	'Experiencer'
				PAT_{N}	$\Leftrightarrow$	'Stimulus'
obava 'fear'	$\equiv$	CPHR_{V}				

(12) *Petr*$_{\text{ACT.nom}}$ *má ze zkoušky*$_{\text{PAT.z+gen}}$ *obavy*$_{\text{CPHR.acc}}$.
'Peter$_{\text{ACT}}$ is afraid of the exam$_{\text{PAT}}$.'

The only exception when light verbs contribute their semantic participant to LVCs is represented by light verbs with a *causative function* – these verbs contribute the semantic participant 'Causator' to LVCs. This participant instigates the events expressed by predicative nouns with which the given light verbs combine. For example, the light verb *poskytovat* 'to provide' has the causative function in the LVC *poskytovat příležitost* 'to provide an opportunity'. The valency frame of this verb is provided in (13). This light verb provides the given LVC with the 'Causator' which is mapped onto its ACT. This 'Causator' serves as an instigator of the situation expressed by the noun *příležitost* 'opportunity', see mapping (16) and example (17).

(13) *poskytovat* 'to provide': ACT_{nom} ADDR_{dat} CPHR_{acc}

(14) *příležitost* 'opportunity': $\text{ACT}_{\text{gen,pos}}$ $\text{PAT}_{\text{gen,k+dat,na+acc,pro+acc,inf,dcc}}$

(15) *naše*$_{\text{ACT.poss}}$ *příležitost zamyslet se*$_{\text{PAT.inf}}$ *nad tím, co je to vina*

(16) *poskytovat příležitost* 'to provide an opportunity':

'Causator'	$\Leftrightarrow$	ACT_{V}					
		ADDR_{V}	$\leftrightarrow$	ACT_{N}	$\Leftrightarrow$	'Agent'	
příležitost 'opportunity'	$\equiv$	CPHR_{V}					
				PAT_{N}	$\Leftrightarrow$	'Future_action'	

(17) *Jeho skutek*$_{\text{ACT.nom}}$ *nám*$_{\text{ADDR.dat}}$ *poskytl příležitost*$_{\text{CPHR.acc}}$ *zamyslet se*$_{\text{PAT.inf}}$ *nad tím, co je to vina.*
'His act$_{\text{ACT}}$ gave us$_{\text{ADDR}}$ an opportunity$_{\text{CPHR}}$ to think$_{\text{PAT}}$ about what is guilt.'

3.1.3. Semantic vs. Syntactic Center of LVCs

The syntactic center of LVCs is represented by the light verb, which can – in contrast to the predicative noun – create a finite clause. However, the semantic core of LVCs is formed by the predicative noun, which contributes its semantic participants to LVCs.

Selecting a particular light verb, the predicative noun can employ its semantic participants in the syntactic structure of a finite clause. Moreover, the choice of a light verb affects the perspective from which the situation expressed by the predicative noun is viewed (see esp. Kettnerová and Lopatková, 2015). Compare, e.g., the LVC *poskytnout dotaci* 'to give a grant' in (18) and the LVC *získat dotaci* 'to obtain a grant' in

(19) where each time a different semantic participant of the noun *dotace* 'grant' occupies the most prominent subject position: the LVC *poskytnout dotaci* 'to give a grant' is perspectivized from the point of view of 'Agent' (*vláda* 'government') whereas the LVC *získat dotaci* 'to obtain a grant' is presented from the perspective of 'Recipient' (*město* 'town'):

(18) *Vláda*$_{\text{ACT}}$ *poskytla městu*$_{\text{ADDR}}$ *dotaci*$_{\text{CPHR}}$ *7 milionů*$_{\text{PAT}}$ *korun.*
'The government$_{\text{ACT}}$ gave a grant$_{\text{CPHR}}$ of 7 million$_{\text{PAT}}$ crowns to the town$_{\text{ADDR}}$.'

(19) *Město*$_{\text{ACT}}$ *získalo od vlády*$_{\text{ORIG}}$ *dotaci*$_{\text{CPHR}}$ *7 milionů*$_{\text{PAT}}$ *korun.*
'The town$_{\text{ACT}}$ obtained a grant$_{\text{CPHR}}$ of 7 million$_{\text{PAT}}$ crowns from the government$_{\text{ORIG}}$.'

3.1.4. Coreference and its Key Role in LVCs

As already mentioned in Section 3.1.2, in LVCs semantically underspecified valency complementations of light verbs obtain their semantic specifications – the principal role in this process is played by coreference between valency complementations of the predicative noun and the light verb.

The most prominent coreferential relation is the one between ACT of a predicative noun with a certain complementation of the light verb. Its prominence is also empirically attested by the corpus material provided by the Prague Dependency Treebank (PDT, Bejček et al., 2013):[7] from 1,695 LVCs with predicative nouns expressed as prepositionless accusative objects of light verbs in PDT, 1,609 LVCs (95% in total) are characterized by the coreference of the nominal ACT and a certain verbal complementation; the remaining 5% represents rather annotation errors (see Kettnerová and Bejček, 2016). The presence of a pair of ACT of the predicative noun and a certain valency complementation of the light verb referring to the same extralinguistic entity can be thus adopted as a definitional criterion for delimiting LVCs. See the test of coreference in Radimský (2010) as well.

Semantically underspecified valency complementations of a light verb can enter into different coreferential relations, depending on valency structure of the predicative nouns selecting the given verb. For example, in the valency frame of the light verb *přinést* 'to bring' in its non-causative function, the ACT and ADDR are semantically unsaturated, see the valency frame of this verb in (20). Three types of coreference of this ACT and ADDR with complementations of predicative nouns are attested in the corpus data, as shown below by simplified dependency trees. Additional two types of coreference appear for the given verb with causative function, when the 'Causator' is mapped onto the verbal ACT. Table 1 below provides more examples of LVCs with the given verb for all the mentioned types of coreference.

I. ACT$_{\text{V}}$ – ACT$_{\text{N}}$ & ADDR$_{\text{V}}$ – ADDR$_{\text{N}}$
First, the ACT of the light verb *přinést* 'to bring' corefers with the nominal ACT and the

[7] http://ufal.mff.cuni.cz/pdt3.0/ (http://hdl.handle.net/11858/00-097C-0000-0023-1AAF-3)

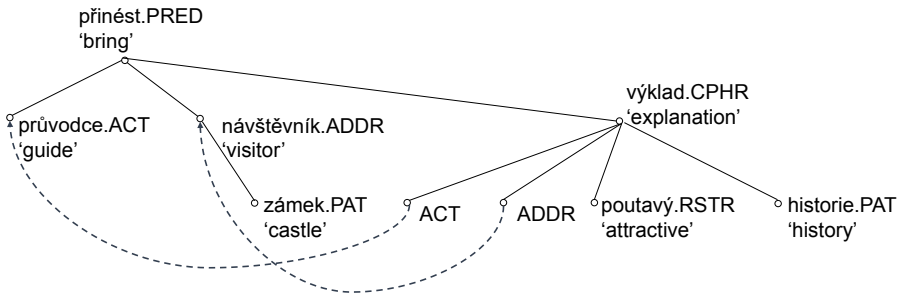

Figure 1. The simplified dependency tree of the LVC 'přinést výklad' 'to bring an explanation' in sentence (24) (the dashed arrows link coreferring valency complementations).

ADDR of the verb corefers with the nominal ADDR. This type of coreference is characteristic, e.g., of the LVC *přinést výklad* 'to bring an explanation'. See the simplified dependency tree of its example sentence (24) in Figure 1. The valency frames of the light verb (20) and of the predicative noun (21) and the scheme of the mapping of semantic participants in the given LVC (23) are provided below.

(20) *přinést* 'to bring': $\mathsf{ACT}_{\mathsf{nom}}\ \mathsf{ADDR}_{\mathsf{dat}}\ \mathsf{CPHR}_{\mathsf{acc}}$

(21) *výklad* 'explanation': $\mathsf{ACT}_{\mathsf{gen,pos,instr,od+gen}}\ \mathsf{ADDR}_{\mathsf{dat}}\ \mathsf{PAT}_{\mathsf{gen,pos,o+loc,dcc}}$

(22) *průvodcův*$_{\mathsf{ACT.poss}}$ *poutavý výklad historie*$_{\mathsf{PAT.gen}}$ *návštěvníkům*$_{\mathsf{ADDR.dat}}$ *zámku*
'guide's$_{\mathsf{ACT}}$ attractive explanation of history$_{\mathsf{PAT}}$ to visitors$_{\mathsf{ADDR}}$ of the castle'

(23) *přinést výklad* 'to bring an explanation':

$$\begin{array}{lcllll} & & \mathsf{ACT_V} & \leftrightarrow & \mathsf{ACT_N} & \Leftrightarrow \text{ 'Speaker'} \\ & & \mathsf{ADDR_V} & \leftrightarrow & \mathsf{ADDR_N} & \Leftrightarrow \text{ 'Recipient'} \\ \textit{výklad} \text{ 'explanation'} & \equiv & \mathsf{CPHR_V} & & & \\ & & & & \mathsf{PAT_N} & \Leftrightarrow \text{ 'Information'} \end{array}$$

(24) *Průvodce*$_{\mathsf{ACT.nom}}$ *přinesl návštěvníkům*$_{\mathsf{ADDR.dat}}$ *zámku poutavý výklad*$_{\mathsf{CPHR.acc}}$ *historie*$_{\mathsf{PAT.gen}}$.
'The guide$_{\mathsf{ACT}}$ brought an attractive explanation$_{\mathsf{CPHR}}$ of history to visitors$_{\mathsf{ADDR}}$ of the castle.'

II. $\mathsf{ACT_V}$ – $\mathsf{ACT_N}$ & $\mathsf{ADDR_V}$ – $\mathsf{PAT_N}$

The second type represents an infrequent type,[8] characterizing, e.g., the LVC *přinést efekt* 'to bring an effect', see the dependency tree in Figure 2 of its example sentence (28).

[8] PAT with predicative nouns in these cases often corresponds to cognitively shifted ADDR.

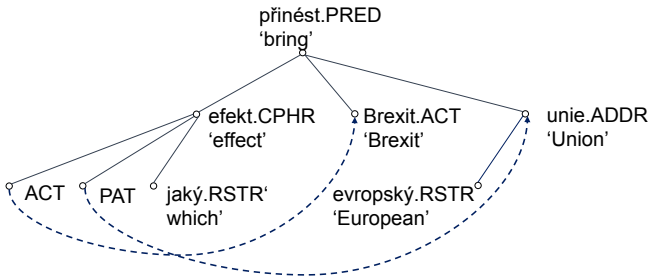

Figure 2. The simplified dependency tree of the LVC 'přinést efekt' 'to bring an effect' in sentence (28).

(25) *efekt* 'effect': $\mathsf{ACT}_{\mathsf{gen,pos}}$ $\mathsf{PAT}_{\mathsf{na+acc,pro+acc}}$

(26) *efekt Brexitu*$_{\mathsf{ACT.gen}}$ *na Evropskou unii*$_{\mathsf{PAT.na+acc}}$
'an effect of Brexit$_{\mathsf{ACT}}$ on the European Union$_{\mathsf{PAT}}$'

(27) *přinést efekt* 'to bring an effect':

$$\begin{array}{lcllcl} & & \mathsf{ACT_V} & \leftrightarrow & \mathsf{ACT_N} & \Leftrightarrow \text{ 'Influencer'} \\ & & \mathsf{ADDR_V} & \leftrightarrow & \mathsf{PAT_N} & \Leftrightarrow \text{ 'Influencee'} \\ \textit{efekt} \text{ 'effect'} & \equiv & \mathsf{CPHR_V} & & & \end{array}$$

(28) *Jaký efekt*$_{\mathsf{CPHR.acc}}$ *přinese Brexit*$_{\mathsf{ACT.nom}}$ *Evropské unii*$_{\mathsf{ADDR.dat}}$*?*
'Which effect$_{\mathsf{CPHR}}$ brings Brexit$_{\mathsf{ACT}}$ to the European Union$_{\mathsf{ADDR}}$?'

III. $\mathbf{ACT_V - PAT_N}$ & $\mathbf{ADDR_V - ACT_N}$

The third type of coreference is typical of, e.g., the LVC *přinést zklamání* 'to bring disappointment', see the dependency tree in Figure 3 of example sentence (32).

(29) *zklamání* 'disappointment': $\mathsf{ACT}_{\mathsf{gen,pos}}$ $\mathsf{PAT}_{\mathsf{nad+instr,z+gen,inf,dcc}}$

(30) *zklamání občanů*$_{\mathsf{ACT.gen}}$ *z výsledků*$_{\mathsf{PAT.z+gen}}$ *voleb*
'disappointment of citizens$_{\mathsf{ACT}}$ by the results$_{\mathsf{PAT}}$ of the elections'

(31) *přinést zklamání* 'to bring disappointment':

$$\begin{array}{lcllcl} & & \mathsf{ACT_V} & \leftrightarrow & \mathsf{PAT_N} & \Leftrightarrow \text{ 'Stimulus'} \\ & & \mathsf{ADDR_V} & \leftrightarrow & \mathsf{ACT_N} & \Leftrightarrow \text{ 'Experiencer'} \\ \textit{zklamání} \text{ 'disappointment'} & \equiv & \mathsf{CPHR_V} & & & \end{array}$$

(32) *Výsledky*$_{\mathsf{ACT.nom}}$ *voleb přinesly občanům*$_{\mathsf{ADDR.dat}}$ *zklamání*$_{\mathsf{CPHR.acc}}$.
'The results$_{\mathsf{ACT}}$ of the elections brought disappointment$_{\mathsf{CPHR}}$ to citizens$_{\mathsf{ADDR}}$.'

Besides its non-causative function, the light verb *přinést* 'to bring' can serve as the causative verb as well (see Section 3.1.2). The 'Causator' provided by this light verb is

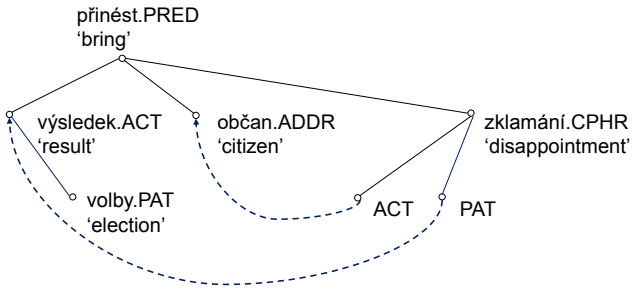

Figure 3. The simplified dependency tree of the LVC 'přinést zklamání' 'to bring disappointment' in sentence (32).

mapped onto its ACT. Thus the only remaining semantically underspecified complementation in the valency frame of the given verb is the ADDR, see the valency frame in (20). LVCs with the causative light verb *přinést* 'to bring' are then characterized by two types of coreference.

IV. 'Causator': $\mathbf{ACT_V}$ & $\mathbf{ADDR_V}$ – $\mathbf{ACT_N}$
This type is characteristic of, e.g., the LVC *přinést poznatek* 'to bring knowledge', see the dependency tree of example sentence (36) in Figure 4.

(33) *poznatek* 'knowledge': $\mathsf{ACT}_{\mathsf{gen,pos}}$ $\mathsf{PAT}_{\mathsf{gen,k+dat,o+loc,dcc}}$ $\mathsf{ORIG}_{\mathsf{z+gen}}$

(34) *naše*$_{\mathsf{ACT.gen}}$ *poznatky o vodě*$_{\mathsf{PAT.o+loc}}$ *na Marsu*
'our$_{\mathsf{ACT}}$ knowledge about water$_{\mathsf{PAT}}$ on Mars'

(35) *přinést poznatek* 'to bring knowledge':

'Causator'	$\Leftrightarrow$	$\mathsf{ACT_V}$				
		$\mathsf{ADDR_V}$	$\leftrightarrow$	$\mathsf{ACT_N}$	$\Leftrightarrow$	'Cognizer'
poznatek 'knowledge'	$\equiv$	$\mathsf{CPHR_V}$				
				$\mathsf{PAT_N}$	$\Leftrightarrow$	'Content'

(36) *Sonda*$_{\mathsf{ACT.nom}}$ *nám*$_{\mathsf{ADDR.dat}}$ *přinesla poznatky*$_{\mathsf{CPHR.acc}}$ *o vodě*$_{\mathsf{PAT.o+loc}}$ *na Marsu.*
'The space probe$_{\mathsf{ACT}}$ brought us$_{\mathsf{ADDR}}$ knowledge$_{\mathsf{CPHR}}$ about water$_{\mathsf{PAT}}$ on Mars.'

V. 'Causator': $\mathbf{ACT_V}$ & $\mathbf{ADDR_V}$ – $\mathbf{PAT_N}$ & $\mathbf{LOC_V}$ – $\mathbf{ACT_N}$
This coreference characterizes, e.g., the LVC *přinést přízeň* 'to bring favor', see the dependency tree in Figure 5 representing example sentence (40).

(37) *přízeň* 'favor': $\mathsf{ACT}_{\mathsf{gen,pos}}$ $\mathsf{PAT}_{\mathsf{dat,k+dat,pro+acc,vůči+dat}}$

(38) *přízeň publika*$_{\mathsf{ACT.gen}}$ *ke zpěvákovi*$_{\mathsf{PAT.k+dat}}$
favor of audience$_{\mathsf{ACT}}$ to the singer$_{\mathsf{PAT}}$

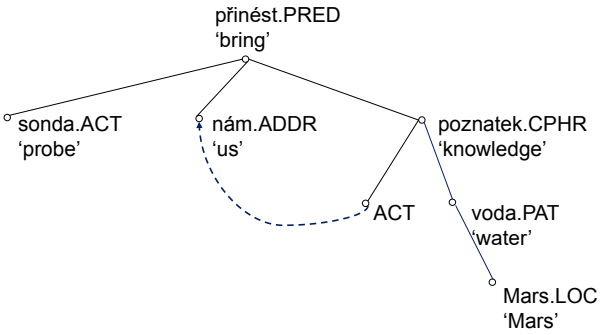

Figure 4. The simplified dependency tree of the LVC ‘přinést poznatek’ ‘to bring knowledge’ in sentence (36).

(39) *přinést přízeň* ‘to bring favor’:

‘Causator’	⇔	ACT_V				
		$ADDR_V$	↔	PAT_N	⇔	‘Stimul’
		LOC_V	↔	ACT_N	⇔	‘Experiencer’
přízeň ‘favor’	≡	$CPHR_V$				

(40) *Zdařilé turné*$_{\text{ACT.nom}}$ *přineslo zpěvákovi*$_{\text{ADDR.dat}}$ *přízeň*$_{\text{CPHR.acc}}$ *u publika*$_{\text{LOC.u+gen}}$.
‘The successful tour$_{\text{ACT}}$ brought the singer$_{\text{ADDR}}$ favor$_{\text{CPHR}}$ of an audience.’

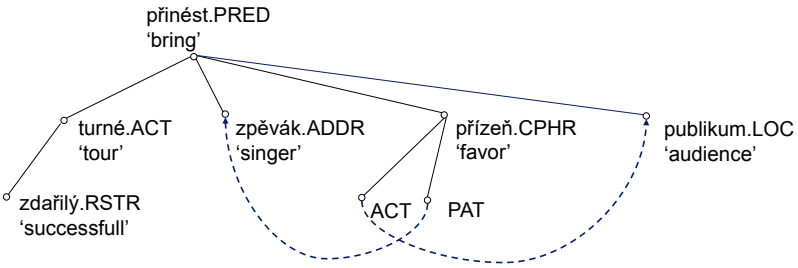

Figure 5. The simplified dependency tree of the LVC ‘přinést přízeň’ ‘to bring favor’ in sentence (40).

Ambiguous mappings

Several light verbs can function ambiguously in a single LVC with respect to non-causative and causative function. For example, the light verb *přinést* ‘to bring’ in the LVC *přinést zklamání* ‘to bring disappointment’ can serve either as non-causative, see above

type III (mapping (31) and example (32)), or as causative one. The latter case belongs to type IV as it exhibits the same coreference as e.g the LVC *přinést poznatek* 'to bring knowledge', see the coreference in (41) and example (42).

(41) *přinést zklamání* 'to bring disappointment':

'Causator'	$\Leftrightarrow$	ACT_V				
		$ADDR_V$	$\leftrightarrow$	ACT_N	$\Leftrightarrow$	'Experiencer'
				PAT_N	$\Leftrightarrow$	'Stimulus'
zklamání 'disappointment'	$\equiv$	$CPHR_V$				

(42) *Radikální dieta*$_{ACT.nom}$ *Janě*$_{ADDR.dat}$ *přinesla jen další zklamání*$_{CPHR.acc}$ *z vlastního selhání*$_{PAT.z+gen}$.
'Radical diet$_{ACT}$ brought Jane$_{ADDR}$ disappointment$_{CPHR}$ by her own failure$_{PAT}$.'

---

type I. $ACT_V - ACT_N$ & $ADDR_V - ADDR_N$
přinést 'to bring': *důkaz* 'evidence', *informace* 'information', *nabídku* 'offer', *návod* 'instruction', *odpověď* 'answer', *podnět* 'impulse', *ponaučení* 'lesson', *výklad* 'explanation', *zprávu* 'message';

type II. $ACT_V - ACT_N$ & $ADDR_V - PAT_N$
přinést 'to bring': *efekt* 'effect';

type III. $ACT_V - PAT_N$ & $ADDR_V - ACT_N$
přinést 'to bring': *neklid* 'uneasiness', *potěšení* 'pleasure', *potíž* 'trouble', *problém* 'trouble', *prospěch* 'benefit', *překvapení* 'surprise', *radost* 'to bring happiness', *uspokojení* 'satisfaction', *užitek* 'benefit', *úleva* 'relief', *útěcha* 'comfort', *zklamání* 'disappointment';

type IV. Causator: ACT_V & $ADDR_V - ACT_N$
přinést 'to bring': *mír* 'calm', *možnost* 'opportunity', *neklid* 'uneasiness', *neštěstí* 'bad luck', *obživa* 'living', *omezení* 'restriction', *pokrok* 'progress', *popularita* 'popularity', *potěšení* 'pleasure', textitpotíž 'trouble', *poznatek* 'insight', *problém* 'trouble', *prospěch* 'benefit', *přátelství* 'friendship', *překvapení* 'surprise', *radost* 'happiness', *riziko* 'risk', *smůla* 'bad luck', *štěstí* 'happiness', *uspokojení* 'satisfaction', *uznání* 'recognition', *užitek* 'benefit', *úleva* 'relief', *úspěch* 'success', *úspora* 'savings', *útěcha* 'comfort', *výhoda* 'advantage', *zisk* 'profit', *zjištění* 'findings', *zklamání* 'disappointment', *zkušenost* 'experience';

type V. Causator: ACT_V & $ADDR_V - PAT_N$ & $LOC_V - ACT_N$
přinést 'to bring': *přízeň* 'favor'.

Table 1. Examples of LVCs with the verb 'přinést' 'to bring' for all relevant types of coreference.

3.2. Surface Syntactic Structure

The surface structure formation of LVCs shares the same basic principle as surface constructions with full verbs according to which each semantic participant is expressed on the surface just once.

When a semantic participant in LVCs corresponds either to a valency complementation of the predicative noun, or to a complementation of the light verb (in case of causative light verbs), the situation is obvious: the surface syntactic expression of such semantic participant can be stipulated only by morphemic forms of the valency complementation (be it nominal, or verbal) to which the given participant corresponds.

However, when a semantic participant is mapped onto a valency complementation of the predicative noun and at the same time (via coreference) onto a complementation of the light verb, it raises a question which of the coreferring valency complementations is expressed on the surface. In these cases, morphemic forms prescribed by individual valency complementations serve as an important clue for the identification of the valency complementations expressed on the surface. As the analysis of extensive corpus data corroborates, these semantic participants are mostly expressed on the surface as valency complementations of the light verb.

For example, let us analyze the surface structure of the LVC *uložit úkol* 'to give a task' in (47): from the semantic participants characterizing this LVC, 'Speaker' and 'Recipient' are mapped onto the ACT and ADDR of the predicative noun and at the same time via coreference onto the ACT and ADDR of the light verb as well, respectively (the scheme of the mapping (46)). 'Speaker' is expressed in the surface structure of the given LVC as the subject corresponding to the verbal ACT as its morphemic form (the nominative case) unequivocally determines, see ACT in the valency frame of the light verb in (43) and ACT in the frame of the noun in (44).

As for 'Recipient', its dative case does not unequivocally indicate the surface position: as both the verbal ADDR and the nominal ADDR can have the form of the dative case, see the valency frames of the light verb (43) and the predicative noun (44), it can be either an indirect object with the function of the verbal ADDR, or an attribute with the function of the nominal ADDR. Thus a question arises which of these complementations the given dative case expresses. The surface syntactic behavior of this complementation in diatheses indicates that it can be accounted for as the verbal ADDR: verbal complementations – in contrast to nominal ones – are sensitive to surface syntactic shifts in diatheses. When recipient passive diathesis is applied to the LVC *uložit úkol* 'to give a task', the given ADDR changes its form from the dative into the nominative, see (48), which clearly manifests that it is the ADDR governed by the light verb. Moreover, the possibility of word order changes (compare (47) and (49)) supports its analysis as verbal ADDR as well.

'Obligation', mapped just onto the PAT of the predicative noun, can be expressed only as the attribute as morphemic forms of the given complementation require, see the valency frame of the noun (44).

(43) *uložit* 'to give': $\mathsf{ACT}_{\mathsf{nom}}\ \mathsf{ADDR}_{\mathsf{dat}}\ \mathsf{CPHR}_{\mathsf{acc}}$

(44) *úkol* 'task': $\mathsf{ACT}_{\mathsf{gen,pos,od+gen}}\ \mathsf{ADDR}_{\mathsf{gen,dat,pos,pro+acc}}\ \mathsf{PAT}_{\mathsf{inf,dcc}}$

(45) *učitelův*$_{\mathsf{ACT.pos}}$ *úkol žákům*$_{\mathsf{ADDR.dat}}$ *narýsovat*$_{\mathsf{PAT.inf}}$ *krychli*
'teacher's$_{\mathsf{ACT}}$ task to students$_{\mathsf{ADDR}}$ to draw$_{\mathsf{PAT}}$ a cube'

(46) *uložit úkol* 'to give a task':

		$\mathsf{ACT}_{\mathsf{V}}$	$\leftrightarrow$	$\mathsf{ACT}_{\mathsf{N}}$	$\Leftrightarrow$	'Speaker'
		$\mathsf{ADDR}_{\mathsf{V}}$	$\leftrightarrow$	$\mathsf{ADDR}_{\mathsf{N}}$	$\Leftrightarrow$	'Recipient'
úkol 'task'	$\equiv$	$\mathsf{CPHR}_{\mathsf{V}}$				
				$\mathsf{PAT}_{\mathsf{N}}$	$\Leftrightarrow$	'Obligation'

(47) *Učitel*$_{\mathsf{ACT.subj.nom}}$ *uložil žákům*$_{\mathsf{ADDR.inobj.dat}}$ *úkol*$_{\mathsf{CPHR.obj.acc}}$ *narýsovat*$_{\mathsf{PAT.attr.inf}}$ *krychli.*
'The teacher$_{\mathsf{ACT}}$ gave students$_{\mathsf{ADDR}}$ a task$_{\mathsf{CPHR}}$ to draw$_{\mathsf{PAT}}$ a cube.'

(48) *Žáci*$_{\mathsf{ADDR.subj.nom}}$ *dostali od učitele*$_{\mathsf{ACT.inobj.od+gen}}$ *uložen úkol*$_{\mathsf{CPHR.obj.acc}}$ *narýsovat*$_{\mathsf{PAT.attr.inf}}$ *krychli.*
'The students$_{\mathsf{ADDR}}$ were assigned a task$_{\mathsf{CPHR}}$ by the teacher$_{\mathsf{ACT}}$ to draw$_{\mathsf{PAT}}$ a cube.'

(49) *Žákům*$_{\mathsf{ADDR.inobj.dat}}$ *učitel*$_{\mathsf{ACT.subj.nom}}$ *uložil úkol*$_{\mathsf{CPHR.obj.acc}}$ *narýsovat*$_{\mathsf{PAT.attr.gen}}$ *krychli.*

3.2.1. Double Expression of a Semantic Participant

There are basically two exceptions from general principles underlying the surface realization of semantic participants. First, in rare cases a semantic participant mapped onto the ACT of a predicative noun and at the same time onto the ACT of a light verb can be expressed twice on the surface. For example, in the LVC *přinést výklad* 'to bring an explanation' in (50), the 'Speaker' is expressed twice in the surface structure, as the subject with the function of the verbal ACT and at the same time as an attribute with the function of the nominal ACT.

(50) *Průvodce*$_{\mathsf{ACT.subj.nom}}$ *přinesl návštěvníkům*$_{\mathsf{ADDR.inobj.dat}}$ *zámku svůj*$_{\mathsf{ACT.attr.pos}}$ *vlastní výklad*$_{\mathsf{CPHR.obj.acc}}$ *historie*$_{\mathsf{PAT.attr.gen}}$.
'The guide$_{\mathsf{ACT}}$ brought his$_{\mathsf{ACT}}$ own explanation$_{\mathsf{CPHR}}$ of history to visitors$_{\mathsf{ADDR}}$ of the castle.'

3.2.2. Semantic Participants Mapped onto an Optional Free Modification of the Light Verb

Second, in those cases in which a semantic participant is mapped onto an actant of the predicative noun and via coreference onto an optional free modification of the light verb, the given semantic participant can be realized either as the nominal valency complementation, or as the verbal one.

For example, in the LVC *probouzet vzpomínku* 'to raise memory', three semantic participants can be expressed on the surface: 'Causator' contributed to the LVC by the causative light verb *probouzet* 'to raise' and two semantic participants – 'Cognizer'

and 'Experience' – provided by the predicative noun *vzpomínka* 'memory', see the valency frame of the given light verb in (51) and the frame of the noun in (52). The LVC *probouzet vzpomínku* 'to raise memory' is characterized by the coreference of the ACT of the predicative noun with the optional LOC of the light verb, see the scheme of the mapping of participants in (54). As a result, this semantic participant can be expressed on the surface either as the verbal LOC (realized on the surface as a local adverbial), see example (55), or as the nominal ACT (expressed on the surface as an attribute), see example (56).

(51) *probouzet* 'to raise': $\mathsf{ACT}_{\mathsf{nom,inf,dcc}}$ $\mathsf{CPHR}_{\mathsf{acc}}$ LOC

(52) *vzpomínka* 'memory': $\mathsf{ACT}_{\mathsf{gen,pos}}$ $\mathsf{PAT}_{\mathsf{na+acc,dcc}}$

(53) *moje*$_{\mathsf{ACT.pos}}$ *vzpomínka na mládí*$_{\mathsf{PAT.na+acc}}$
'my$_{\mathsf{ACT}}$ memory of (my) youth$_{\mathsf{PAT}}$'

(54) *probouzet vzpomínku* 'to raise memory':

'Causator'	$\Leftrightarrow$	$\mathsf{ACT}_{\mathsf{V}}$				
vzpomínka 'memory'	$\equiv$	$\mathsf{CPHR}_{\mathsf{V}}$				
		$\mathsf{LOC}_{\mathsf{V}}$	$\leftrightarrow$	$\mathsf{ACT}_{\mathsf{N}}$	$\Leftrightarrow$	'Cognizer'
				$\mathsf{PAT}_{\mathsf{N}}$	$\Leftrightarrow$	'Experience'

(55) *Hudba*$_{\mathsf{ACT.subj.nom}}$ *ve mně*$_{\mathsf{LOC.adv.v+loc}}$ *probouzela vzpomínky*$_{\mathsf{CPHR.obj.acc}}$ *na mládí*$_{\mathsf{PAT.attr.na+acc}}$.
music$_{\mathsf{ACT}}$ in me$_{\mathsf{LOC}}$ raised memories$_{\mathsf{CPHR}}$ on youth$_{\mathsf{PAT}}$

(56) *Hudba*$_{\mathsf{ACT.subj.nom}}$ *probouzela moje*$_{\mathsf{ACT.attr.pos}}$ *vzpomínky*$_{\mathsf{CPHR.obj.acc}}$ *na mládí*$_{\mathsf{PAT.attr.na+acc}}$.
music$_{\mathsf{ACT}}$ raised my$_{\mathsf{ACT}}$ memories$_{\mathsf{CPHR}}$ on youth$_{\mathsf{PAT}}$
'The music raised my memories of my youth.'

3.2.3. Principles Governing the Expression of Semantic Participants

The principles governing the expression of semantic participants in the surface structure of LVCs can be summarized as follows:[9]

- A semantic participant mapped onto a single valency complementation is expressed in a surface syntactic position:
  - of the respective valency complementation of the predicative noun (be it actant, or free modification),
  - of the respective valency complementation of the light verb (representing 'Causator').
- A semantic participant mapped onto two coreferring valency complementations is expressed in a surface syntactic position:
  - either of the respective valency complementation of the light verb (be it actant, or free modification),

[9] As the double expression of a semantic participant in LVCs discussed above in Section 3.2.1 is highly restricted, it is not reflected in the given principles.

– or of the respective valency complementation of the predicative noun if the coreferring verbal complementation is an optional free modification.

4. Formal Model of Lexicographic Representation of LVCs

The lexicographic representation of LVCs requires a close cooperation of both the lexical and the grammar part of the language description. In this section, we describe how the information on LVCs is reflected in the data and grammar component of the VALLEX lexicon, see above Section 2. As we have demonstrated above, the deep and surface syntactic structure of LVCs can be inferred from valency structures of both light verbs and predicative nouns and coreferential relations between individual valency complementations on the rule basis. These rules, described in the grammar component of the lexicon, operate on the information provided by its data component.

4.1. Data Component

In the data component, three special attributes have been introduced providing the information necessary for deriving deep and surface syntactic structure of LVCs.

Attribute `lvc`. The data component of the VALLEX lexicon stores lexical units of both predicative nouns and light verbs,[10] providing their core valency characteristics in the form of valency frames, as described in Section 3.1.1 and 3.1.2. The respective lexical units are interlinked via the special attribute `lvc`: with each lexical unit of a predicative noun, this attribute provides references to lexical units of the light verbs selected by the given noun, whereas with each lexical unit of a light verb, links to the predicative nouns with which the light verb combines are provided.

Attribute `map`. This attribute is provided within lexical units of light verbs. It stores the information on coreference between valency complementations of light verbs and complementations of predicative nouns – the coreference is represented as a set of pair(s) of coreferring valency complementations, the valency complementations being represented by their respective functors (and for user's convenience indicated in the lower index either as `V`, or as `N`, distinguishing verbal complementations from nominal ones).

Attribute `instig`. This attribute is provided within lexical units of causative light verbs – it gives the valency complementation from the respective valency frames onto

[10] Although collocations of predicative nouns with light verbs form multiword lexical units, we use here the term lexical units to refer to individual predicative nouns and light verbs, stressing their compositional possibilities. We are aware that in case of light verbs, this term is not, strictly speaking, correct: light verbs are not able to occur outside the collocation with predicative nouns (giving evidence of their semi-lexical status). However, from the formal point of view, light verbs can be treated similarly as individual senses of full verbs.

which the semantic participant 'Causator' is mapped, the given complementation being represented by its respective functor.

Each valency frame of a light verb is assigned with a set of pair(s) or triplet(s) (if the attribute `instig` is relevant) of the above given attributes, distinguishing different coreference relations and eventually causative function of the light verb; relevant pairs (or triplets) are differentiated by Arabic numerals. In addition, each set is accompanied with examples illustrating individual LVCs. See the lexical entry of the light verb *uložit* 'to impose' in Figure 8 and the entry of the predicative noun *úkol* 'task' in Figure 7.

4.2. Grammar Component

In the grammar component, formal rules governing both the deep and surface structure formation of LVCs are stored, operating on the information provided by the data component, as follows:

The *deep syntactic structure of LVCs* consists of:

(i) all valency complementations of the light verb provided by its valency frame (attribute `frame`),
(ii) all valency complementations of the predicative noun provided by its valency frame (attribute `frame`),
(iii) coreferential relations between individual valency complementations of the light verb and complementations of the predicative noun (attribute `map`).

The *surface syntactic structure of LVCs* comprises:[11]

A. syntactic positions of all valency complementations of the light verb, namely:
   (i) the syntactic position of the predicative noun
   (its morphemic form is given by the `CPHR` valency complementation),
   (ii) the syntactic position of 'Causator', if relevant
   (its morphemic form is given by the valency complementation provided in the attribute `instig`),
   (iii) the syntactic position(s) of other valency complementations which corefer with complementations of the predicative noun (only optional free modifications may remain unexpressed)
   (morphemic forms of these positions are given by the respective valency complementations provided in the attribute `map`).
B. syntactic positions of those valency complementations of the predicative noun that satisfy the following conditions:

[11] These principles summarize the surface structure formation of LVCs on which further surface syntactic operations leading to deletion of valency complementations can be applied due to their optionality, actual ellipsis, generalization etc.

(iv) the syntactic position(s) of the valency complementations that do not corefer with any complementations of the light verb
(morphemic forms of these positions are given by the respective valency complementations of the predicative noun),
(v) the syntactic position(s) of the valency complementations that corefer with optional free modifications of the light verbs not expressed on the surface
(morphemic forms of these positions are given by the respective valency complementation of the predicative noun).[12]

5. Annotation of LVCs in VALLEX

5.1. Lexical Stock

Each LVC is formed by a collocation consisting of a predicative noun and a light verb. The large amount of such collocations in Czech is not easily manageable, thus, some selection criteria had to be determined at the beginning of the annotation process. For the identification of the collocations representing LVCs, we first had to identify an inventory of verb lemmas already stored in the VALLEX lexicon that can function as light verbs. For this purpose, we used the valency lexicon PDT-Vallex (Urešová, 2011). In this lexicon, valency frames of light verbs are indicated by the functor CPHR, labeling the valency position of predicative nouns. On this basis, we automatically identified 124 verb lemmas that have at least one valency frame in PDT-Vallex with the CPHR functor. The intersection of verb lemmas obtained from PDT-Vallex and of verb lemmas contained in VALLEX was 105 in total. As VALLEX treats aspectual counterparts of verbs as a single verb lexeme, respective aspectual counterparts have been added (if not already in the list). The resulting number of 145 verb lemmas has formed the inventory of verbs selected for the further annotation.

To identify the most frequent and most salient predicative nouns which select the given light verbs, we used the Sketch Engine (Kilgarriff et al., 2014), a corpus tool allowing users to obtain summaries of words' grammatical and collocational properties. A balanced corpus of synchronous texts SYN2010[13] was used as the material base. In the first step, the collocation lists of all selected verb lemmas were obtained. From these lists, only the nominal collocates expressed as direct object in the accusative case (representing the central and most frequent cases of light verb collocations in Czech) were selected: almost 3,050 noun lemmas (21 nouns on average for a verb lemma); see also Kettnerová et al. (2016).

[12] In rare cases, the syntactic position of the nominal ACT coreferring with the verbal ACT is expressed in the surface structure (typically as a possessive pronoun), in addition to its expression in the verbal position according to principle (Aiii), compare Section 3.2.1 and example (50).

[13] Czech National Corpus – SYN2010. Institute of the Czech National Corpus, Faculty of Arts, Charles University, Prague 2010. Available from http://www.korpus.cz.

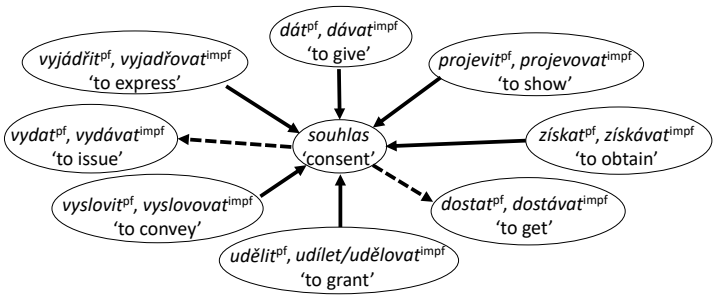

Figure 6. Collocations of the predicative noun 'souhlas' 'consent' with light verbs, found with the Sketch Engine; the collocations marked by the continuous arrow have been identified in the lists of nominal collocates of the given light verbs 'dát, dávat' 'to give', 'projevit, projevovat' 'to show' etc.; the collocation joined by the dashed arrow have been extracted from the list of verbal collocates of the noun 'souhlas' 'consent').

A human annotator has been asked to indicate only those nouns in each list that represent predicative nouns forming collocations with the given light verb. As the main criterion for distinguishing collocates with light verbs from those with full verbs, the coreference of the nominal ACT with a valency complementation of the light verb has been adopted (see Section 3.1.4). The collocations that satisfy this condition are interpreted as LVCs.

In the second step, salient collocations of the predicative nouns have been added as well, using the Sketch Engine: for each of the predicative noun obtained in the first step, its missing relevant collocations with light verbs have been automatically extracted from its verb collocation list (in this step, we restricted the verbs to the list of 145 already identified light verb lemmas). See Figure 6, displaying collocations of the noun *souhlas* 'consent' obtained in the first and second steps of the annotation process.

The resulting number of collocations from the first and second step sums up to 2,991 collocations in total (counted as combinations of a lemma of a light verb and a lemma of a predicative noun). These collocations represent the lexical stock integrated into the VALLEX lexicon.

5.2. Annotation Process

In the next step, the selected 2,991 collocations of predicative nouns with light verbs have been assigned with the relevant information, as introduced in Section 4. Each verb lemma and noun lemma have been processed separately, and then interlinked into a relevant LVC.

5.2.1. Annotation of Predicative Nouns

As the VALLEX lexicon originally contained the information only on valency behavior of verbs, it was necessary to add the same information on the selected predicative nouns. The data format of the lexicon, as described in Section 2, is suitable also for the description of other parts-of-speech. We thus made use of the same structure of lexical entries as designed for verbs (excluding verb-specific optional attributes).

The lexical entry of a predicative noun comprises its noun lexeme, i.e., a two-fold unit associating all forms of the given noun with its lexical units (its individual senses). Each lexeme is represented by the respective noun lemma of the predicative noun, or (if relevant) by more than one lemma.

úkol

1 ≈ zadání; skutečnost určená k vykonání `assignment'
-frame: **ACT$_{2,pos,od+2}$ ADDR$_{2,3,pos,pro+4}$ PAT$_{inf,dcc}$**
-lu: úkol-1
-lvc: **ukládat-uložit-8**, zadat-zadávat-4, klást-4, přebírat-přebrat-převzít-2.1, uvalit-uvalovat-2
-example: Pracovní úkol zaměstnavatele.ACT vyplývá zejména z popisu práce …;
V tomto směru jsou úkoly od vedení.ACT klubu celkem jasné.;
Hráči přesně plní trenérovy.ACT úkoly.;
Hosté prohrávali a jejich.ADDR úkolem od trenéra.ACT bylo neprohrát utkání o deset branek.;
domácí úkoly pro předškoláky.ADDR;
těžký úkol vybrat.PAT vítěze.;
Výsledkem byl úkol Komisi.ADDR, aby začala.PAT připravovat pro tři země akční programy.;
Nový čínský úkol zní: ať naši zpěváci dobudou.PAT svět.;
Z rady vyvstal úkol, že bychom měli svolat.PAT obecnou diskuzi.
-control: ADDR

2 ≈ poslání; úloha; funkce `mission; role; function'
-frame: **ACT$_{2,pos}$ PAT$_{inf,dcc}$**
-lu: úkol-2
-lvc: vykonat-vykonávat-2, přebírat-přebrat-převzít-2.2
-example: Úkolem řešitele.ACT sudoku je v co nejkratším čase doplnit prázdná místa v tabulce.;
Úkolem závodníka.ACT bude předvést odvážný improvizovaný skok do vody.;
Jaký je náš.ACT úkol v dějinách člověčenstva?;
Před zápasem byl můj.ACT úkol určit.PAT brankářku, která nastoupí do zápasu.;
Je to jeho.ACT úkol, aby zakázka byla.PAT čistá a hotová včas.;
To je taky jeho.ACT hlavní úkol, ať sešívky zapomenou.PAT na poháry a boje o místo nahoře.;
Jeho.ACT předpokládaný úkol, že má sedět.PAT při každém přelíčení a podat.PAT čtenářům jeho průběh, byl nesmírně obtížný.
-control: ACT

Figure 7. Lexical entry of the predicative noun 'úkol' 'task' (two lexical units).

The key information on the valency behavior of individual lexical units of the predicative noun is provided by valency frames, exemplified by illustrative examples. Morphemic forms in valency frames of the noun describe the usage of its lexical units in nominal structures, see Section 3.1.1.

For the purpose of the description of LVCs, each lexical unit of a predicative noun is assigned with the special attribute `lvc` providing a list of references to individual light verbs with which the given noun forms LVCs. See the illustrative example of the lexical entry of the predicative noun *úkol* 'task' in Figure 7.

We have restricted the number of syntactically annotated predicative nouns to those that form LVCs in the data with at least two light verbs (277 noun lexemes represented by 284 noun lemmas). Those nouns that represent a part of LVCs with one light verb are provided only in the form of a list of noun lemmas in the attribute `lvc` in the lexical entries of the respective light verbs (if these noun lemmas are counted, the number of predicative nouns increases to 577 lemmas). The basic statistics on annotated predicative nouns is provided in Table 2.

	Lemmas	Lexemes	Lexical Units
Predicative nouns	284 (577)	277	350
Light verbs	145	78	117

Table 2. The basic statistics on the annotated predicative nouns and the light verbs in VALLEX. The number of noun lemmas in parenthesis includes also predicative nouns which form LVCs with only one light verb.

5.2.2. Annotation of Light Verbs

In the lexical entry of each verb lemma indicated in the previous step as a lemma representing a light verb, a relevant lexical unit (or more lexical units) of the given light verb has been identified. If no relevant lexical unit has been comprised in the lexical entry, it has been manually added. Then each identified lexical unit has been subject to necessary adjustment.

First, in each valency frame of relevant lexical units, a valency complementation standing for predicative nouns has been labeled with the CPHR functor. Second, each lexical unit of the light verb has been ascribed with one or more sets of the following attributes (as introduced in Section 4): (i) the attribute `lvc`, providing references to lexical units of the predicative nouns with which the given light verb form LVCs, and (ii) the attribute `map`, introducing the information on coreference characterizing the given LVCs. (iii) In case of causative light verbs, each set is further supplemented with the attribute `instig`, storing the information on the valency position of 'Causator'. If more

ukládat^{impf} , uložit^{pf}

....

8 ≈ light verb `to give; to assign' (as light verb)
-frame: **ACT$_1$ ADDR$_3$ CPHR$_4$**
-lu: ukládat-uložit-8
-lvc1: **úkol-1**, zákaz-1
-map1: ACTv-ACTn, ADDRv-ADDRn
-example1:
impf: Prostřednictvím ředitele úřadu ministr a náměstci … ukládají úkoly úředníkům.;
Pokyn ministerstva obrany ukládal složkám resortu přísný zákaz ničit po výcviku munici.
pf: Vladimir Putin uložil vládě úkol něco se situací dělat už před osmi lety.;
… řidičce městský úřad uložil zákaz činnosti spočívající v zákazu řízení motorových vozidel.
....
-lvc3: omezení-1, omezení-2, opatření-1, pokání-1, povinnost-1
-map3: ADDRv-ACTn
-instig3: ACT
-example3:
impf: Nová technická pravidla ukládají odstranění či omezení některých počítačových systémů.;
Současná společnost ukládá člověku početná omezení.;
ČNB se musí striktně držet zákonných postupů a může ukládat opatření k nápravě.;
… také počítač provádí hříšníka úvodními modlitbami, ptá se na jeho hříchy a ukládá pokání.;
Soud zpravidla ukládá radnicím povinnost zajistit nájemníkům náhradní bydlení.
pf: Také ruská státní agentura Rosselkhoznadzor … si vymínila právo uložit omezení dovozu, …;
Soud může pachateli uložit omezení, aby se zdržel řízení.;
Proto jsme také městským službám neuložili žádné zvláštní opatření.;
Zřejmě to ale nebyl jen tak ledaskdo, protože sám papež uložil Bernardu Ignácovi pokání.;
Soud uložil žalované společnosti povinnost zaplatit kanceláři zmíněnou částku s příslušenstvím.

Figure 8. Simplified lexical unit of the light verb 'uložit' 'to give'.

than one set of these attributes is relevant with a single lexical unit of a light verb, the sets are distinguished by Arabic numerals. (iv) In addition, each LVC is exemplified by corpus example provided in the attribute `example`, attached to individual sets of the above given attributes. See the example of the lexical unit of the light verb *uložit* 'to give' in Figure 8.

As a result, 117 lexical units characterizing light verbs have been annotated; these lexical units are contained in 78 verb lexemes, represented by 145 verb lemmas, see Table 2. Most light verbs are non-causative: from 117 lexical units in total, 89 underlie non-causative light verbs in the annotated data and 28 lexical units correspond to causative light verbs. The semantic participant 'Causator' has been mapped predominantly onto ACT of light verbs, in less cases it has corresponded to ORIG. For the basic statistics on the mapping of 'Causator' see Table 3.

'Causator'	Lexical Units	LVCs
ACT	25	325 (365)
ORIG	3	13 (13)

Table 3. The basic statistics on the mapping of 'Causator' with causative light verbs in VALLEX; the first column provides the number of lexical units representing causative light verbs, the second one gives the number of LVCs in which light verbs are of causative character (counted as combinations of individual lexical units of light verbs and lexical units of predicative nouns). The number in parenthesis includes also LVCs with nouns that combine only with one light verb.

5.3. Types of Coreference

As a result of the annotation, almost 1,500 different LVCs, counted as combinations of individual lexical units of light verbs and lexical units of predicative nouns, have been identified. Each LVC is characterized by a certain type of coreference between verbal and nominal valency complementations (or by more than one type), see Section 3.1.4. The information on coreference has been assigned to the valency frame of a light verb (attribute `map`). For causative light verbs, also the information on 'Causator' mapping is provided (attribute `instig`).

In the annotated data, most LVCs are characterized by a single type of coreference and (if relevant) by the mapping of 'Causator'; however, almost 200 annotated LVCs allow for different types of coreference, often distinguished by the presence/absence of 'Causator'.

In the annotation, 21 different types of coreference (distinguished with respect to the presence of 'Causator') have been identified for all LVCs (including 7 ambiguous types). The most frequent type of coreference is represented by the coreference of the ACT of a non-causative light verb with the ACT of a predicative noun; this type can be exemplified by, e.g., the LVC *mít obavu* 'to be afraid; lit. to have fear' (11). Within the group of causative light verbs, the coreference of the verbal ADDR and the nominal ACT with the mapping of 'Causator' onto the verbal ACT is the most frequent one; it characterizes, e.g., the LVC *přinést poznatek* 'to bring knowledge' (35). Table 4 provides all types of coreference identified in the annotated data, illustrated by examples.

6. Conclusion

In this paper, we have summarized results of a theoretical analysis of syntactic behavior of Czech light verb constructions. We have focused both on their deep and surface syntactic structure, demonstrating the syntactic compositionality of these constructions. We have deepened an insight into a key role of coreference between valency complementations of the light verb and the predicative noun forming an LVC

Type 1		Type 2		Number	Example
'Causator'	Coreference	'Causator'	Coreference		
	ACT_V–ACT_N			716 (902)	*udělat chybu* 'to make a mistake'
	ACT_V–ACT_N, $ADDR_V$–$ADDR_N$			182 (200)	*činit výtku* 'make a reproach'
	ACT_V–PAT_N, LOC_V–ACT_N	ACT_V	LOC_V–ACT_N	141 (151)	*vyvolat strach* 'to raise fear'
	ACT_V–ACT_N, $ADDR_V$–PAT_N			99 (142)	*klást odpor* 'put up resistance'
ACT_V	$ADDR_V$–ACT_N			68 (82)	*poskytnout útěchu* 'to give comfort'
	ACT_V–$ADDR_N$, $ORIG_V$–ACT_N			60 (67)	*přijmout poděkování* 'to accept thanks'
ACT_V	LOC_V–ACT_N			60 (65)	*vzbouzet myšlenky* 'to provoke thoughts'
	ACT_V–PAT_N, $ORIG_V$–ACT_N			35 (44)	*dostat péči* 'to get care'
	ACT_V–PAT_N, $ADDR_V$–ACT_N	ACT_V	$ADDR_V$–ACT_N	35 (44)	*činit potěšení* 'to make pleasure'
	ACT_V–PAT_N, LOC_V–ACT_N			30 (37)	*nacházet oporu* 'to find support'
	ACT_V–ACT_N, LOC_V–PAT_N			13 (17)	*nalézt zalíbení* 'to develop a taste'
$ORIG_V$	ACT_V–ACT_N			13 (13)	*získat výhodu* 'to gain an advantage'
ACT_V	BEN_V–ACT_N			8 (10)	*otevřít přístup* 'to open an access'
	ACT_V–ACT_N		ACT_V–PAT_N, LOC_V–ACT_N	8 (8)	*ztratit důvěru* 'to lose confidence'
	ACT_V–$ORIG_N$, LOC_V–ACT_N			7 (7)	*budit dojem* 'to give an impression'
	ACT_V–$ADDR_N$, LOC_V–ACT_N			7 (7)	*udělat zkoušku* 'to pass an exam'
	ACT_V–$ORIG_N$, LOC_V–ACT_N	ACT_V	LOC_V–ACT_N	6 (6)	*probouzet pocit* 'to inspire a feeling'
	ACT_V–$ADDR_N$, LOC_V–ACT_N	ACT_V	LOC_V–ACT_N	4 (4)	*vzbudit podezření* 'to raise suspicion'
ACT_V	$ADDR_V$–PAT_N, LOC_V–ACT_N			2 (2)	*přinést přízeň* 'to bring favor'
	ACT_V–$ADDR_N$, LOC_V–ACT_N		ACT_V–PAT_N, LOC_V–ACT_N	1 (1)	*budit soustrast* 'to arouse sympathy'
	ACT_V–ACT_N	ACT_V	BEN_V–ACT_N	1 (1)	*vytvářet zisk* 'to make a profit'

Table 4. The basic statistics on the coreference identified with the annotated LVCs (individual combinations of lexical units of light verbs and lexical units of predicative nouns), taking into account the type of coreference and the 'Causator' mapping (if relevant). If more than one type is characteristic of a single LVC, these types are distinguished as type 1 and type 2. The number in column 5 gives the number of LVCs in VALLEX; in parenthesis, the number including also those LVCs with nouns combining with a single light verb follows.

in the process of the syntactic formation of the given LVC, emphasizing a role of mapping of semantic participants characterizing the given LVC onto valency complementations.

The proposed theoretical analysis has been verified (and refined) within a linguistic annotation of a large amount of light verb constructions. The resulting description has been used in the VALLEX lexicon for their theoretically adequate and economic representation. The data have been published – after both manual and automatic data consistency checking –- in the new version of the VALLEX lexicon, release 3.5.[14]

The VALLEX 3.5 lexicon comprises annotation of almost 3,000 collocations of predicative nouns with light verbs (counted as combinations of a lemma of a light verb and a lemma of a predicative noun), which correspond to almost 1,500 LVCs (counted as individual combinations of a lexical unit of a light verb and a lexical unit of a predicative noun). The LVC annotation has affected almost 350 newly created lexical units of predicative nouns and 120 lexical units of light verbs. The syntactic formation of LVCs has been described in the grammar component of VALLEX in a form of syntactic rules operating on the information from the data component (namely valency frames of respective light verbs and predicative nouns and three special attributes describing the LVC formation `lvc`, `map`, and `instig`).

Acknowledgements

The work on this project has been supported by the grant of the *Czech Science Foundation* (project GA15-09979S); the data annotation has been partially supported by the LINDAT/CLARIN project of the *Ministry of Education, Youth and Sports of the Czech Republic* (project number LM2015071).

This work has been using language resources distributed by the LINDAT/CLARIN project of the *Ministry of Education, Youth and Sports of the Czech Republic* (project number LM2015071).

Bibliography

Alonso Ramos, Margarita. Towards the Synthesis of Support Verb Constructions: Distribution of Syntactic Actants between the Verb and the Noun. In Wanner, L. and I. A. Mel'čuk, editors, *Selected Lexical and Grammatical Issues in the Meaning-Text Theory*, pages 97–137. John Benjamins Publishing Company, Amsterdam, Philadelphia, 2007.

Bejček, Eduard, Eva Hajičová, Jan Hajič, Pavlína Jínová, Václava Kettnerová, Veronika Kolářová, Marie Mikulová, Jiří Mírovský, Anna Nedoluzhko, Jarmila Panevová, Lucie Poláková, Magda Ševčíková, Jan Štěpánek, and Šárka Zikánová. *Prague Dependency Treebank 3.0*. Univerzita Karlova v Praze, MFF, ÚFAL, Prague, Czech Republic, 2013. `http://ufal.mff.cuni.cz/pdt3.0`.

[14] `http://ufal.mff.cuni.cz/vallex/3.5/`

Butt, Miriam. The Light Verb Jungle: Still Hacking Away. In Amberber, Mengistu, Brett Baker, and Mark Harvey, editors, *Complex Predicates in Cross-Linguistic Perspective*, pages 48–78. Cambridge University Press, Cambridge, 2010.

Butt, Miriam and Wilhelm Geuder. *On the (semi)lexical status of light verbs*, volume 59 of *Studies in Generative Grammar*, pages 323–370. Mouton de Gruyter, Berlin – New York, 2001.

Grimshaw, Jane and Armin Mester. Light verbs and θ-marking. *Linguistic inquiry*, 19(2):205–232, 1988.

Hinrichs, Erhard, Andreas Kathol, and Tsuneko Nakazawa. *Complex Predicates in Nonderivational Syntax. Syntax and Semantics 30*. Academic Press, San Diego, 1998.

Jespersen, Otto. *A Modern English Grammar on Historical Principles, Part VI, Morphology*. Allen and Unwin, 1965.

Kettnerová, Václava. Syntaktická struktura komplexních predikátů. *Slovo a slovesnost*, 18(1): 3–24, 2017.

Kettnerová, Václava and Eduard Bejček. Distribution of Valency Complements in Czech Complex Predicates: Between Verb and Noun. In Calzolari, Nicoletta, Khalid Choukri, Thierry Declerck, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Asunción Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the 10th International Conference on Language Resources and Evaluation (LREC 2016)*, pages 515–521, Paris, France, 2016. European Language Resources Association. ISBN 978-2-9517408-9-1.

Kettnerová, Václava and Markéta Lopatková. The Representation of Czech Light Verb Constructions in a Valency Lexicon. In Hajičová, Eva, Kim Gerdes, and Leo Wanner, editors, *Proceedings of the Second International Conference on Dependency Linguistics, Depling 2013*, pages 147–156, Praha, Czechia, 2013. Univerzita Karlova v Praze, Matfyzpress. ISBN 978-80-7378-240-5.

Kettnerová, Václava and Markéta Lopatková. At the Lexicon-Grammar Interface: The Case of Complex Predicates in the Functional Generative Description. In Hajičová, Eva and Joakim Nivre, editors, *Proceedings of Depling 2015*, pages 191–200, Uppsala, Sweden, 2015. Uppsala University.

Kettnerová, Václava and Markéta Lopatková. Complex Predicates with Light Verbs in VALLEX: From Formal Model to Lexicographic Description. In Hlaváčová, Jaroslava, editor, *Proceedings of the 17th Conference on Information Technologies - Applications and Theory (ITAT 2017)*, volume 1885 of *CEUR Workshop Proceedings*, pages 15–22, Praha, Czechia, 2017a. ÚFAL MFF UK, CreateSpace Independent Publishing Platform.

Kettnerová, Václava and Markéta Lopatková. Ke koreferenci u komplexních predikátů s kategoriálním slovesem. *Korpus – gramatika – axiologie*, 16:3–26, 2017b.

Kettnerová, Václava, Petra Barančíková, and Markéta Lopatková. Lexicographic Description of Complex Predicates in Czech: Between Lexicon and Grammar. In Margalitadze, Tinatin and George Meladze, editors, *Proceedings of the XVII EURALEX International Congress: Lexicography and Linguistic Diversity*, Tbilisi, Georgia, 2016. Ivane Javakhishvili Tbilisi State University, Tbilisi University Press.

Kettnerová, Václava, Veronika Kolářová, and Anna Vernerová. Deverbal Nouns in Czech Light Verb Constructions. In Mitkov, Ruslan, editor, *Lecture Notes in Artificial Intelligence, Computational and Corpus-Based Phraseology. Second International Conference, Europhras 2017. London, UK, November 13???14, 2017.*, volume 10596 of *Lecture Notes in Computer Science*, pages 205–219, Cham, Switzerland, 2017. European Association for Phraseology EUROPHRAS, University of Wolverhampton, Association for Computational Linguistics, Bulgaria, Springer.

Kilgarriff, Adam, Vít Baisa, Jan Bušta, Miloš Jakubíček, Vojtěch Kovář, Jan Michelfeit, Pavel Rychlý, and Vít Suchomel. The Sketch Engine: ten years on. *Lexicography ASIALEX*, 1(1): 7–36, 2014.

Lopatková, Markéta and Jarmila Panevová. Recent Developments in the Theory of Valency in the Light of the Prague Dependency Treebank. In Šimková, Mária, editor, *Insight into Slovak and Czech Corpus Linguistics*, pages 83–92. Veda, Bratislava, 2006.

Lopatková, Markéta, Václava Kettnerová, Eduard Bejček, Anna Vernerová, and Zdeněk Žabokrtský. *Valenční slovník českých sloves VALLEX*. Karolinum, Praha, 2016.

Panevová, Jarmila. On Verbal Frames in Functional Generative Description I–II. *The Prague Bulletin of Mathematical Linguistics*, 22–23:3–40, s. 17–52, 1974–75.

Panevová, Jarmila. *Formy a funkce ve stavbě české věty*. Academia, Praha, 1980.

Panevová, Jarmila. Valency Frames and the Meaning of the Sentence. In Luelsdorff, Philip A., editor, *The Prague School of Structural and Functional Linguistics*, pages 223–243. John Benjamins Publishing Company, Amsterdam/Philadelphia, 1994.

Radimský, Jan. *Verbo-nominální predikát s kategoriálním slovesem*. Editio Universitatis Bohemiae Meridionalis, České Budějovice, 2010.

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

Urešová, Zdeňka. *Valence sloves v Pražském závislostním korpusu*. Ústav formální a aplikované lingvistiky, Praha, 2011.

Address for correspondence:
Václava Kettnerová
kettnerova@ufal.mff.cuni.cz
Institute of Formal and Applied Linguistics
Faculty of Mathematics and Physics,
Charles University
Malostranské náměstí 25
118 00 Praha 1, Czech Republic