Natural Language Processing, Corpus Linguistics, E-learning

Seventh International Conference
Bratislava, Slovakia, 13–15 November 2013
Proceedings

Editors
Katarína Gajdošová
Adriána Žáková

RAM-Verlag
2013

The articles have been reviewed by members of the Program Committee.

Slovak National Corpus
Ľ. Štúr Institute of Linguistics
Slovak Academy of Sciences
Bratislava, Slovakia 2013
`http://korpus.juls.savba.sk/~slovko/`

Corpus Based Identification of Czech Light Verbs

Václava Kettnerová[1], Markéta Lopatková[1], Eduard Bejček[1],
Anna Vernerová[1], and Marie Podobová[2]

[1] Faculty of Mathematics and Physics, Charles University in Prague, Czech Republic

[2] Institute of Slavonic Studies, The Academy of Sciences of the Czech Republic,
Prague, Czech Republic

Abstract. In this paper, we describe a corpus based experiment focused on the possibility of identifying Czech light verbs. The experiment had two main aims: (i) to establish the inventory of Czech light verbs entering into combinations with predicative nouns, and (ii) to verify the adopted criteria for distinguishing light usages from full usages of the given verbs. As for the inventory of light verbs, we propose and verify the hypothesis that the possibility of light usages of verbs is related to their semantic class membership rather than to their high frequency. In the second part of the experiment, we exploited the compiled inventory of Czech verbs inclined to occur as light verbs. The criteria adopted for distinguishing light usages of a verb from its full ones were applied to selected corpus sentences with these verbs. Three annotators were asked to determine whether a verb occurrence in an extracted corpus sentence corresponds to the full or to the light usage of the given verb. The feasibility of this task has been proven by the achieved κ_w statistics 0.686 and by the inter-annotator agreement 85.3%. As a result of this experiment, we obtained 893 verbonominal combinations of Czech light verbs and predicative nouns. These combinations will be further utilized for the lexicographic representation of these phenomena.

1 Introduction

As a verb represents the most important syntactic unit in a language, the description of its syntactic structure belongs to the primary tasks of both theoretical and computational linguistics. At present, the basic information on syntactic behavior of verbs is provided in many lexical resources. However, the description of advanced syntactic properties of verbs is still missing. Light verbs belong to such advanced linguistic phenomena. In this case, the syntactic structure of a sentence is not solely determined by a verb alone but also by a predicative noun with which the verb combines. Predicative nouns exhibit two characteristic properties: (i) they denote an event meaning, e.g. a process, a state, or a property, not a resultative meaning, and (ii) they are characterized by a set of valency complementations. See the following examples with the verb *nést* ‘to carry’ (1)–(2). Unlike the syntactic structure with the full usage of the verb in (1), the syntactic structure with the light usage of the verb (2) is affected also by the predicative noun *odpovědnost* ‘responsibility’. This predicative noun expresses ‘the state of being responsible’ and contributes its valency complementation *za bezpečnost* ‘for security’ to the resulting syntactic structure.

(1) *Učitel nese sešity.*
'The teacher is carrying exercise books.'

(2) *Učitel nese odpovědnost za bezpečnost žáků.*
the teacher – carries – responsibility – for the security of pupils
'The teacher is responsible for the security of (his) pupils.'

While easily mastered by native speakers, light verbs pose a serious challenge for foreign speakers as well as for automated language processing (esp. for machine translation, information extraction, information retrieval, question answering, etc.) [11]. It has been recognized for a long time that both language learning and NLP tasks would be facilitated by a lexical resource providing an explicit and systematic representation of these phenomena. Such a representation should be based on a thorough theoretical analysis of light verbs.

Despite being subject to many analyses, many aspects of light verbs are still not clear. Even a clearcut definition of light verbs as a specific group of verbs is lacking [6]. In accordance with [2], we consider light verbs as a specific usage of a verb that loses individual semantic properties and retains only some of semantic facets of its full verb counterpart. To acquire semantic capacity, a light verb combines with a predicative noun which contributes its individual semantic features to a resulting complex predicate. From this point of view, each verb can potentially be used as a full or a light verb.

In this paper, we report on a corpus based analysis of Czech light verbs that enter into combinations with predicative nouns. Considering the wide range of issues related to light verbs, this analysis focuses on the possibility to identify them. Such an identification requires operational criteria for distinguishing light usages of a verb from its full usages, see examples (1)–(2). The criteria adopted here were verified in the parallel annotation of a large amount of corpus data obtained from the Czech National Corpus.[1] The annotation process and the criteria used in the annotation are described in detail in Section 3.

At the beginning of the annotation, lemmas of the verbs that are prone to combine with predicative nouns were necessary to select. Compiling the inventory of such verb lemmas is described in Section 2. In order to draw up this inventory, the valency lexicons VALLEX[2] and PDT-VALLEX[3] were explored. Let us briefly introduce these two valency lexicons.

1.1 Lexical Resources

Both VALLEX and PDT-VALLEX take the Functional Generative Description (henceforth FGD, [12]) as their theoretical background, and both are human as well as machine readable.

VALLEX is a valency lexicon of Czech verbs (see esp. [14], [8]) providing rich syntactic information on roughly 2,730 lexemes containing 6,460 lexical units ('senses'). Unlike traditional dictionaries, VALLEX treats a pair of perfective and imperfective aspectual counterparts as a single lexeme – if perfective and imperfective verbs were counted separately, the size of the lexicon would virtually grow to 4,250 verb entries. Almost one half

[1] `http://ucnk.ff.cuni.cz/`

[2] `http://ufal.mff.cuni.cz/2.5/`

[3] `http://ufal.mff.cuni.cz/lindat/PDT-Vallex.html/`

of the lexical units are sorted into 22 rough semantic groups (e.g. verbs of communication, motion, transport, exchange, mental action). At present, this lexicon inconsistently describes several randomly selected light verb usages: most light verb usages are subsumed under valency frames corresponding to full verb usages, less of them are represented by separate valency frames. In both cases, an explicit indication of the light verb usage is missing. A preliminary theoretical analysis of an interaction between verbal and nominal valency structures within light verb constructions and the possibility of its adequate description in the VALLEX lexicon is provided in [7].

PDT-VALLEX is a valency lexicon which was built during the annotation of the Prague Dependency Treebank[4] (henceforth PDT [4]) as a supporting annotation tool designed for preserving data consistency in the annotated corpus, see esp. [5], [13]. It describes valency behavior of Czech verbs, nouns, adjectives and adverbs in a fully formalized way (7,500 verbs, 3,800 nouns, 800 adjectives, and a few adverbs). Only those senses of words that occurr in the annotated data of PDT or some other treebanks in the PDT family (Prague Czech-English Dependency Treebank[5] and Prague Dependency Treebank of Spoken Language)[6] are recorded. The information on syntactic structures with light verbs is simply listed in pairs of unlinked separate valency frames: (a) in valency frames of light verbs, in which the valency complementation occupied by a predicative noun is labeled with the CPHR functor (for CompoundPHRaseme, see [4]), and (b) in valency frames of predicative nouns.

2 Inventory of Czech Light Verbs

Although light verbs and their full verb counterparts always have identical forms, they only agree in some of their semantic aspects [2], [9]. For instance, the light verb *nést* ‘to carry’ in (2) loses the individual meaning of its full counterpart in (1): the meaning of the full verb can be characterized as ‘the shared movement of a physical object and a person who controls the shared path’, but in the light usage in (2), the lexical properties expressing the physical movement are suppressed; instead, the lexical features of ‘permanent presence of a certain sense experienced by a person wherever they move’ are foregrounded. The individual lexical semantic property ‘a sense of moral commitment or obligation to somebody or something’ is supplied by the predicative noun *odpovědnost* ‘responsibility’ with which the verb combines.

The question arises whether the verbs that are inclined to combine with predicative nouns share some common features on the basis of which they can be characterized. According to [3], high frequency verbs, occurring in various semantic and syntactic contexts, have a strong tendency to lose their individual semantic properties and to combine with predicative nouns. The hypothesis was formulated on Swedish verbs.

However, Czech verbs do not confirm this hypothesis: we sorted Czech verbs according to their frequency in the Czech National Corpus[7] (henceforth CNC); modal verbs and the verb *být* ‘to be’ (with primarily auxiliary function) were excluded. We compared the obtained list of the first 50 Czech verbs with the verbs with the CPHR functor in PDT-VALLEX, i.e., those verbs that are classified as used as light verbs in at least one of their

[4] `http://ufal.mff.cuni.cz/pdt2.0/`

[5] `http://ufal.mff.cuni.cz/pcedt2.0/`

[6] `http://ufal.mff.cuni.cz/pdtsl/cz/`

[7] The balanced subcorpus of contemporary Czech texts SYN2000 was used.

occurrences in PDT. Only 32% of the high frequency verbs contain at least one valency frame with the CPHR functor in PDT-VALLEX.

Thus we tried to establish the criteria for the identification of Czech verbs with possible light usages on another basis. As a starting point, we carried out a tentative survey of the verbs with the CPHR functor in PDT-VALLEX.[8] This study revealed an interesting fact: verbs with valency frame(s) describing light usages fall into just a few semantic groups. They express exchange (e.g. *vzít* 'to take', *dát* 'to give'), location (e.g. *pokládat, položit* 'to put down'), motion (e.g. *přicházet, přijít* 'to come'), transport (e.g. *vést* 'to lead'), or they refer to an action in a generic way (e.g. *dělat* 'to do').

In the next step, we further explored the idea that the capacity of a verb to be used as light is related to their semantic class membership. We sorted all verbs belonging to one of the above mentioned semantic groups in VALLEX according to their frequency in CNC.[9] The verbs designating actions in a generic way are not grouped together in a specific semantic class in this lexicon. However, as they represent a significant group of verbs allowing for light usages, we have included all 17 verb lemmas with generic meaning obtained from PDT-VALLEX directly into our experiment. The most frequent verbs of exchange, location, motion, and transport, plus the verbs with generic meaning were chosen as candidates for light verbs. From the list of candidates, verbs with less than six valency frames in VALLEX were removed. In order to achieve a satisfactory coverage of verbs in CNC, we selected first 59 most frequent verb lemmas – this number covers 20.0% of total verb occurrences in CNC.

The resulting inventory of selected verbs was exploited in the second part of the experiment focusing on the possibility of distinguishing light usages from full usages. This strategy to identify Czech lemmas allowing for light usages gave more satisfactory results than frequency – 48 from the overall 59 selected verbs (81.4%) were used as light verbs, based on the findings of our experiment (Section 3).

Alternatively, there was an option to directly use the verbs with the CPHR functor from PDT-VALLEX as candidates for light verbs. This method would eliminate verb lemmas predicating only as full verbs; however, many verbs that can form a light usage would be missed: 27 out of 48 verbs with a light usage that occurred in the annotation do not have the CPHR functor in PDT-VALLEX.

3 Annotation of Light Verbs

In this section, we describe in detail the annotation of corpus sentences based on 59 selected Czech verb lemmas. For each of these selected verb lemmas included in the experiment, 100 random sample sentences were extracted from the CNC. Thus the annotated data size is 5,900 sentences per each annotation.

Three human annotators in parallel were asked to determine whether a verb occurrence in an extracted sentence corresponds to a full or a light usage of the given verb (thus the

[8] PDT-VALLEX contains 148 Czech verb lemmas with at least one valency frame containg the CPHR functor.

[9] Although phase verbs are usually considered to represent light verbs, they are not indicated by the CPHR functor in PDT-VALLEX. As a result, these verbs were included in our experiment only if they fall into some of the above mentioned semantic groups in VALLEX. For instance, the verb *skončit* 'to finish' was included in the experiment as it is classified both as a phase verb and as a verb of location according to VALLEX.

overall number of annotated sentences is 17,700). The main aim of this annotation was to examine the native speakers' agreement on the interpretation of light verb usages. To facilitate the interpretation, the annotators could take a context of one preceding sentence into consideration. When the annotators indicated that a given occurrence of a verb is a light usage, they had to determine the whole verbonominal combination of the given light verb and a predicative noun. Also, an uncertainty flag indicating that the annotators are not quite sure could be attached to a positive answer.

3.1 Criteria for Distinguishing Light Verb Usages from Full Ones

At the beginning of the annotation, it was necessary to single out criteria for distinguishing light verb usages from their full usages. For this purpose, we have adopted two criteria mentioned in the rich bibliography on light verbs – the reduction test and the test of coreference of nominal and verbal complementations.

Reduction Test. The reduction test, proposed in [10], is based on the assumption that it is the predicative noun (not the light verb) that represents the semantic core of the entire verbonominal combination. As a result, it is the predicative noun that stands for the whole verbonominal combination and it cannot be omitted from the combination – in contrast to the light verb. This test consists of the sequence of two syntactic operations: (i) relativization and (ii) omission of the light verb. When these operations are applied on a particular syntactic structure – e.g. on the sentence structure in (2) repeated here as (3), (i) the relativization results in (4) and (ii) the omission results in (5) – the semantic invariant between (4) and (5) is clearly preserved. However, when this test is applied to a full verb usage – e.g. on (1) repeated here as (6), (i) the relativization results in (7) and (ii) the omission results in (8), the semantic invariant is not preserved: (8) is obviously not equivalent to (7).

(3) *Učitel nese odpovědnost za bezpečnost žáků.*
the teacher – carries – responsibility – for the security of pupils
'The teacher is responsible for the security of pupils.'

(4) *Odpovědnost, kterou za bezpečnost žáků nese učitel.*
the responsibility – which – for the security of pupils – carries – the teacher

(5) *Odpovědnost učitele / Učitelova odpovědnost za bezpečnost žáků.*
'The teacher's responsibility for the security of pupils.'

(6) *Učitel* nese sešity.
'The teacher carries exercise books.'

(7) *Sešity, které nese učitel.*
'Exercise books, which the teacher carries.

(8) *Sešity učitele / Učitelovy sešity.*
'Teacher's exercise books.'

Coreference Test. The second criterion stipulates that some of valency complementations of a light verb and a predicative noun within the resulting complex predication must be referentially identical. This condition is imposed esp. on the ACTor of the event expressed by a predicative noun. For example, in the verbonominal combination of the light verb *udělat* ‘to make’ and the predicative noun *dojem* ‘impression’, resulting in the verbonominal combination *udělat dojem* ‘to make an impression’, the ACTor of the predicative noun *dojem* corefers with the ACTor of the verb *udělat*, see Figure 1. Apart from the ACTor, other valency complementations of a predicative noun can be coreferentially related to verbal complementations of a light verb. For example, in the combination *dostat příkaz* ‘to get an order’, two valency complementations from the nominal valency frame ACTor and ADDResse corefer with the verbal valency complementations ORIGin and ACTor, respectively, see Figure 2.

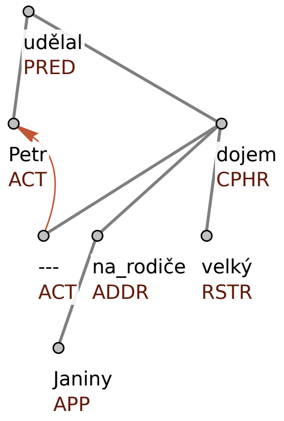

Fig. 1. The (simplified) dependency tree for the sentence *Petr udělal na Janiny rodiče velký dojem* ‘Peter made a great impression on Jane’s parents’

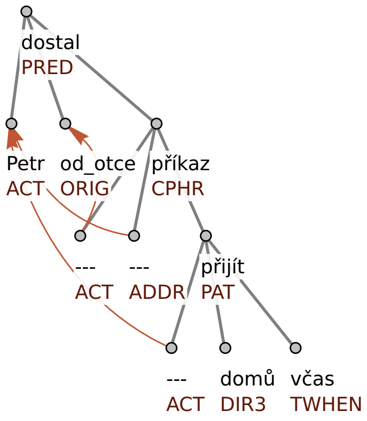

Fig. 2. The (simplified) dependency tree for the sentence *Petr dostal od otce příkaz přijít domů včas* ‘Peter was given the order from his father to come home on time’

Supporting Criteria. In addition to the reduction test and the coreference test, the annotators could rely on other auxiliary criteria. These criteria are based on the observation that the possibility for predicative nouns to be pronominalized (9)-(10) or to be asked for by wh-questions (9)-(11) is highly restricted in verbonominal constructions with a light verb.

(9) *Petr upadl do rozpaků.*
Peter – fell – in embarrassment

(10) **Petr upadl do toho.*
Peter – fell – in this

(11) **Do čeho Petr upadl?*
in what – Peter – fell?

3.2 Annotation Task

In case that the annotators concluded that a light usage of a verb is present in a given sentence (also a special flag indicating an uncertainty of light usage could be used), they had two tasks. First, they had to indicate the whole combination of the light verb and the predicative noun. In case that a sentence contained coordinated predicative nouns combined with a single light verb, the annotators had to determine all of the predicative nouns combined with the given light verb. Second, after identifying a light usage of a verb, the annotator had to determine with which verbal valency complementation the valency complemenation 'ACTor' of the given predicative noun is coreferential.

The annotated data size and overall statistics on the annotations are summarized in Tables 1 and 2.

Annotated verbs	59
Annotated sentences for each verb	100
Parallel annotations	3
Total annotated sentences	17,700

Table 1. Annotated data size

Annotator:	A	B	C
Verb lemmas with light usages	40	35	49
Verb lemmas with uncertain light usages	6	13	4
Light verb usages	713	522	699
Uncertain light verb usages	110	256	287
Full verb usages	5,077	5,122	4,914
Total verb usages	5,900	5,900	5,900
Found verbonominal combinations	843	796	1,002
Coreference between nominal 'ACTor' and some of verbal complementation(s)	615	649	755

Table 2. Overall statistics on the annotations

3.3 Inter-Annotator Agreement

Table 3 provides inter-annotator agreement (IAA), Cohen's κ and Artstein and Poesio's α_κ statistics on the annotated 17,700 sentences.

First two columns represent IAA, i.e., the percentage of sentences with an agreement out of all annotated sentences (the first column); if also the combinations "yes-maybe" is considered to be an agreement (the second column), the pairwise inter-annotator agreement naturally rises.

The first row introduces the average pairwise agreement. In our case of three parallel annotations, it is the sum of agreements of three possible annotation pairs divided by three. The second row shows an IAA for all three annotations together, which is a more rigid measure than the pairwise average: e.g. combinations "yes-yes-maybe" and "yes-maybe-no" are considered as disagreements for an exact match (left column), thus obtaining 0 and 0, respectively (whereas corresponding pairwise average values in the first row would be $\frac{1+0+0}{3} = \frac{1}{3}$ and $\frac{0+0+0}{3} = 0$, respectively). The latter combination "yes-maybe-no" is a disagreement (=0) even with uncertainty tolerance (but would be rated $\frac{1+0+0}{3} = \frac{1}{3}$ in the average pairwise match).

The third column represents Cohen's κ (generalized for multiple annotators in the last row), i.e., the inter-annotator agreement above the chance counted from individual annotators' preferences for their answers (roughly speaking, from their 'average answers'). The last column is a weighted variant of κ (called κ_w for two annotators and α_κ for multiple annotators, see an extended version of [1]). Weights for disagreement were set as follows: "yes-no" is not an agreement at all ($a_{\text{yes,no}} = 0$), but "yes-maybe" and "no-maybe" is counted as a partial agreement ($a_{\text{yes,maybe}} = \frac{2}{3}$ and $a_{\text{no,maybe}} = \frac{1}{3}$). The third row shows generalizations of κ coefficients for multiple annotators, which is claimed to be "a better practise" than an average of pairs by [1].

	IAA exact match	IAA uncertainty tolerance	κ unweighted	κ_w, α_κ weighted
Average pairwise match	89.7%	92.6%	0.602	0.688
Match of all three annotators	85.3%	88.9%		
Agreement above chance for three annotators			0.600	0.686

Table 3. Inter-annotator agreement and κ statistics of three parallel annotations

3.4 Golden Data

The sentences with exact agreement across three involved annotations form the so called golden data. The sentences with disagreement were manually re-annotated in order to resolve disagreement and to unify the annotations. On the basis of the golden data, we obtained verbonominal combinations which can be further applied in the lexicographic description of Czech light verbs. The overall statistics on the golden data is provided in Table 4. The numbers of light usages of each verb involved in the experiment are provided in the Appendix.

Annotated sentences	5,900
Sentences with light verb usages	855
Sentences with uncertain light verb usages	18
Sentences with full verb usages	5,027
Verbonominal combinations	893
Verb lemmas annotatated	59
Verb lemmas with light usages	48

Table 4. Overall statistics on the golden data

4 Conclusion

We have described an experiment with the identification of Czech light verb usages. This experiment consisted of two parts: in the first part, we have explored the possibility to identify the inventory of Czech verbs allowing for light usages; in the second part, we have examined the criteria adopted for distinguishing light usages of a verb from its full ones.

As a result of the first part, we have suggested the hypothesis that the possibility of Czech verbs to be used as light verbs is connected to their semantic class membership (exchange, location, motion, transport verbs and verbs with generic meaning) rather than to their high frequency. However, the hypothesis has to be further examined esp. on low frequency verbs belonging to the selected semantic groups.

In the second part of the experiment, the reliability of distinguishing light usages of a verb from its full ones on the basis of the adopted criteria – the reduction test and the coreference test – has been examined. The achieved inter-annotator agreement (IAA 85.3% and κ_w 0.686) appears to be promising.

Further, as a result of the annotation process, the golden data that consists of the sentences with exact agreement and the sentences with a resolved disagreement has been obtained. The golden data contains 893 instances of combinations of light verbs and predicative nouns. These data will be further exploited as the lexical stock in the lexicographic representation of Czech light verbs in the valency lexicon of Czech verbs VALLEX.

Acknowledgments

The research reported in this paper has been supported by the Czech Science Foundation GA ČR, grant No. GA P406/12/0557 and partially grant No. GA P406/10/0875. This work has been using language resources stored and/or distributed by the LINDAT-Clarin project of MŠMT (project LM2010013).

References

[1] Artstein, R. and Poesio, M. (2008). Inter-Coder Agreement for Computational Linguistics. *Computational Linguistics*, 34(4):555–596. Extended version is available at: `http://cswww.essex.ac.uk/Research/nle/arrau/`.

[2] Butt, M. (2010). The Light Verb Jungle: Still Hacking Away. In Mengistu Amberber, B. B. and Harvey, M., editors, *Complex Predicates in Cross-Linguistic Perspective*, pages 48–78. Cambridge University Press, Cambridge.

[3] Cinková, S. (2009). *Words that Matter: Towards a Swedish-Czech Colligational Dictionary of Basic Verbs*, volume 2 of *Studies in Computational and Theoretical Linguistics*. Institute of Formal and Applied Linguistics, Prague.

[4] Hajič, J., Hajičová, E., Panevová, J., Sgall, P., Pajas, P., Štěpánek, J., Havelka, J., and Mikulová, M. (2006). *Prague Dependency Treebank 2.0*. Linguistic Data Consortium, Philadelphia, PA, USA.

[5] Hajič, J., Panevová, J., Urešová, Z., Bémová, A., Kolářová, V., and Pajas, P. (2003). PDT-VALLEX: Creating a Large-Coverage Valency Lexicon for Treebank Annotation. In *Proceedings of The Second Workshop on Treebanks and Linguistic Theories*, volume 9, pages 57–68.

[6] Hanks, P., Urbschat, A., and Gehweiler, E. (2006). German Light Verb Constructions in Corpora and Dictionaries. *International Journal of Lexicography*, 19(4):439–457.

[7] Kettnerová, V. and Lopatková, M. (2013). The Representation of Czech Light Verb Constructions in a Valency Lexicon. In *Proceedings of the Dependency Linguistics Conference, DepLing 2013*. (accepted).

[8] Lopatková, M., Žabokrtský, Z., and Kettnerová, V. (2008). *Valenční slovník českých sloves*. Nakladatelství Karolinum, Praha.

[9] Macháčková, E. (1979). *Analytická spojení typu sloveso + abstraktní substantivum (analytické vyjadřování predikátů)*. Ústav pro jazyk český ČSAV, Praha.

[10] Radimský, J. (2010). *Verbo-nominální predikát s kategoriálním slovesem*. Editio Universitatis Bohemiae Meridionalis, České Budějovice.

[11] Sag, I. A., Baldwin, T., Bond, F., Copestake, A., and Flickinger, D. (2002). Multiword Expressions: A Pain in the Neck for NLP. In *Proceedings of the Third International Conference on Intelligent Text Processing and Computational Linguistics (CICLING 2002)*, pages 1–15, Mexico City, Mexico.

[12] Sgall, P., Hajičová, E., and Panevová, J. (1986). *The Meaning of the Sentence in Its Semantic and Pragmatic Aspects*. Reidel, Dordrecht.

[13] Urešová, Z. (2011). *Valence sloves v Pražském závislostním korpusu*, volume 8 of *Studies in Computational and Theoretical Linguistics*. Institute of Formal and Applied Linguistics, Prague.

[14] Žabokrtský, Z. and Lopatková, M. (2007). Valency Information in VALLEX 2.0: Logical Structure of the Lexicon. *The Prague Bulletin of Mathematical Linguistics*, (87):41–60.

Appendix: List of verbs annotated within the experiment

This appendix lists the verbs involved in the annotation process sorted with decreasing frequency in CNC. It provides the numbers of their light usages stored in the golden data. Some of these verbs were not found capable of creating light usages (at least in the examined data) – these have a dash in the second column.

verb	light usages	*uncertain* light usages
mít	36	1
jít	7	
stát	3	1
dát	23	1
dostat	30	2
uvést	12	
přijít	13	
dělat	14	
získat	43	
patřit	–	
vést	38	
udělat	26	1
najít	3	
zůstat	3	
dojít	35	
vrátit	2	
nechat	4	
platit	–	
vzít	6	
dosáhnout	52	1
skončit	64	
připravit	9	1
ukázat	1	
přijmout	26	
vydat	15	
počítat	–	
vypadat	–	
vyjít	2	1
chodit	–	
postavit	4	1
sedět	–	
držet	4	3
ztratit	43	1
uzavřít	34	1
ležet	2	
měnit	–	
vybrat	–	
zastavit	9	
podat	19	
pohybovat	–	
jezdit	–	
založit	3	
nést	28	
končit	50	1
přidat	4	
projít	7	
obrátit	4	
učinit	30	
vystoupit	2	
stavět	4	1
sejít	2	
udržet	22	1
položit	15	
padnout	27	
převzít	37	
pustit	20	
hodit	–	
přejít	17	
zvednout	1	