Semantics of noun/verb conversion in Czech: Lessons learned from corpus data annotation

Magda Ševčíková, Hana Hledíková, Lukáš Kyjánek, Anna Staňková
Charles University, Prague, Czech Republic

The paper contributes to the study of conversion in Czech, which in the literature on Czech is rarely understood as an independent word-formation process and therefore not much attention has been paid to its semantics (only as part of research on derivation). An annotation experiment is presented in which semantic relations in 800 conversion pairs are assessed on the basis of corpus concordances by two human raters in parallel. The annotation experiment provides evidence of the applicability of semantic classification of conversion based on meaning paraphrases on large-scale linguistic data, showing that it is possible to reach reasonable inter-annotator agreement. The sources of disagreement, as well as the patterns of polysemy found in the data, provide information about the conceptual relationships between the semantic categories in conversion.

Keywords: *conversion, semantics, Czech, noun, verb*

1 Introduction

Conversion has received more attention in English than in other languages, partly because the concept is driven by the properties of English and it is less clear how to handle it in languages with different morphological features. The English-based definition of conversion as the transfer of a word into a different word class without a change of its form (cf. Adams 1973; Bauer 1983; Štekauer 1996; Plag 2003; Bauer et al. 2013, among others) has been relaxed when examining comparable word formation in other languages. Valera (2015) also discusses conversion without word-class change, but with formal identity of the input and output, conversion as word-class change without formal identity, and conversion without word-class change and without formal identity in the context of European languages. Formal changes that are involved in the latter types of conversion are related to the addition or removal of inflectional material. All the types of conversion discussed can thus be brought under the definition of conversion as word formation without derivational material (cf. Štekauer et al. 2012).

In conversion between nouns and verbs in Czech, which is the focus of the present paper, the input and output do not have identical citation forms because the verbal inflections are not shared by the nouns; cf. the thematic suffix (conveying the imperfective grammatical aspect) and the infinitival ending which are dropped when the verb *řezat* 'to cut' converts to the noun in (1), and the addition of these inflectional markers to the noun in the reverse process of forming the verb *olejovat* 'to oil' in (2).[1]

---

[1] In the examples and in the experiment presented here, the direction of conversion is inferred from how the verb from the given conversion pair forms its aspectual counterpart. As has been demonstrated for Czech (Ševčíková 2021a, 2021b), the aspect-changing strategy of verbs in Czech correlates with the direction in noun/verb conversion as determined by semantic criteria: Those verbs that are considered the result of verb-to-noun conversion form the aspectual counterpart by substituting the thematic suffix for a different one; cf. the noun *řez* 'cut' in (1) with its presumably source (imperfective) verb *řez-a-t* 'to cut' whose perfective counterpart *říz-nou-t* 'to cut' is formed by replacing the theme *-a-* with *-nou-*. In contrast, verbs that are semantically based on nouns (cf. *olejovat* 'to oil' in (2)) do not combine with another theme and a prefix can be attached to them to make the aspect change.

(1)   *řez-a-t*   >   *řez*
       cut-IPFV-INF   cut
       'to cut'   'cut'

(2)   *olej-ova-t*   <   *olej*
       oil-IPFV-INF   oil
       'to oil'   'oil'

However, the involvement of mandatory inflectional markers on verbs, which distinguishes non-prototypical conversion in Czech and other Slavic languages (cf. Cetnarowska 1996 or Manova 2011) from central conversion in English, does not reduce the range of semantic changes that are conveyed by conversion (cf. Hledíková 2022) and the description of these changes is as much a challenge in Czech as in English.

　　The main goal of this paper is to apply semantic categories to large-scale linguistic data (800 conversion pairs extracted from a corpus of written Czech) and to examine whether reasonable agreement can be reached when the same data is evaluated by multiple raters. Also, based on previous research, we assume that the semantic categories are not equally distributed, i.e. that some appear in conversion more frequently than others, and that each semantic category does not have the same relationship to each other semantic category – some meanings may be more closely conceptually related than others. We use the results of the annotation experiment to investigate the relationships between the semantic categories as they are manifested in frequent inter-annotator disagreements and in patterns of multiple categories which appear together in a single conversion pair in the final disambiguated annotation.

　　This paper starts with a brief review of existing semantic classifications of conversion and their application in linguistic research (Section 2). Section 3 introduces the semantic categories that are used in our experiment, and the compilation of the dataset to which the categories are assigned by two annotators in parallel. In Section 4, the parallel annotation is evaluated and sources of disagreement are analyzed. As part of this analysis, we also resolve cases with divergent annotations, yielding final data in which we observe the co-occurrence of several semantic categories in individual conversion pairs (polysemy). The paper concludes with a short summary in Section 5.

2 Related work

Although the understanding of conversion in Czech has been confronted with the concept in English (cf. Dokulil 1968a, 1968b; 1982), the account of conversion as an independent word-formation process is rather rare in Czech linguistics (cf. Bednaříková 2009 or Bozděchová 2016). In the mainstream literature on Czech, this type of formation is considered as a type of derivation and also the meaning of converted words is discussed together with words with overt derivational affixes (Dokulil 1962; Dokulil et al. 1986; Štícha 2018, among others).

　　In this section, we review selected classifications that have been developed specifically to describe semantic change during conversion. Most of these focus on English, but concepts that include another language or deal with a different language than English are also included. Some authors provide semantic categories for both directions (noun to verb and verb to noun), some focus only on one of them. The authors use different criteria to delimit the individual semantic categories.

2.1 Semantic classification of conversion based on paraphrase sentences

An influential line of semantic research consists in approaches that examine the relationship between the noun and the verb in a conversion pair on the basis of the syntactic structure of the converted word's meaning paraphrase.

Marchand (1969) uses the syntactic structure of the paraphrase which describes the meaning of the resulting noun/verb to classify conversion pairs into eight syntactic-semantic relations (four for the denominal direction, four for the deverbal direction). For instance, the verb *to father* is classified as the Predicate–Subject Complement relation because the noun *father* is the subject complement in the verb's paraphrase "to be the father". Adams' (1973) classification is similar, as it is also based on the grammatical function that the motivating word has in the meaning paraphrase of the motivated word. In cases where the noun has the role of an adverbial, semantic subtypes are delimited, namely INSTRUMENTAL (e.g. *to hammer*), LOCATIVE (*to pocket*), and "to remove N from somebody/something" (*to bone*).

Clark & Clark's (1979) classification is also based on a syntactic analysis of paraphrase sentences, but goes deeper than the previous proposals in the sense that it takes into consideration what function (case, semantic role) the motivating noun has in the argument structure of the paraphrase. The authors postulate nine categories for denominal converted verbs: AGENT (e.g. *to butcher*), DURATION (*to summer*), EXPERIENCER (*to witness*), GOAL (*to powder*), INSTRUMENT (*to bicycle*), LOCATION (*to kennel*), LOCATUM (*to blanket*), SOURCE (*to piece*), and MISCELLANEOUS verbs (*to lunch*).

Plag (1999) also works with meaning paraphrases and classifies denominal converted verbs into ten classes, such as INSTRUMENTAL ("to use N"), LOCATIVE ("to put into N"), or ORNATIVE ("to provide with N"). Bauer et al. (2013) adopt Plag's categories for denominal conversion and add categories for the deverbal direction; e.g. AGENT ("people or person who V-s"), INSTANCE ("an instance of V-*ing*"), or INSTRUMENT ("the thing that verbs, a way of V-ing"). Many of the categories overlap with Cetnarowska's (1993) classification of deverbal converted nouns, which works with similar meaning paraphrases.

Gottfurcht (2008) applies Plag's classification on a sample of denominal verbs (including converted verbs) from the *Oxford English Dictionary* (OED). Each semantic category is assigned its lexical conceptual structure (e.g. RESULTATIVE: CAUSE [x BE [noun base]], ORNATIVE: CAUSE [GO [noun base] TO y]) and the annotation is carried out by mapping the verbs' sense definitions from the OED onto one of the lexical conceptual structures.

Valera (2020, 2023) uses semantic labels based mainly on Plag to annotate a sample of English and Spanish denominal verbs extracted from the *British National Corpus* and from a comparable corpus of Spanish, respectively. The analysis is carried out on individual concordances, i.e. the verb is annotated for the meaning that it has in a particular context. The semantic label assignment is based on dictionary records. The analysis revealed a large number of cases where different concordances of the same lemma are classified into different semantic categories.

2.2 Classifications based on other criteria

Štekauer (1996) approaches conversion as a conceptual recategorization of objects of the extra-linguistic reality. The objects are classed in one of four onomasiological categories (SUBSTANCE, ACTION, QUALITY, and CONCOMITANT CIRCUMSTANCE). The semantic classification of conversion is then described in terms of onomasiological connectives, or

semantic roles, which express the relation between the conceptual categories of the input and of the output item in conversion. For instance, the relation between the SUBSTANCE category of the motivating naming unit *display* and the ACTION category of the motivated unit *to display* is conveyed by the connective INSTRUMENT OF ACTION. Ten semantic roles are applied in the analysis of new English conversions (besides INSTRUMENT OF ACTION, e.g. PLACE OF ACTION, CAUSE OF ACTION, A HUMAN BEING PERFORMING THE ACTION; Štekauer 1996: 135–140).

Semantic classifications of conversion are also provided in works which treat conversion as metonymy. Martsa (2013) provides semantic categories based on the metonymic relation (i.e. a conceptual mapping from one entity to another within one domain/schema/frame) which underlies the semantic relation between the noun and the verb in the conversion pair. For instance, the meaning of the verb *to ski* is described as the INSTRUMENT FOR ACTION metonymy in relation to the motivating noun *ski*. The author also points out that not only metonymy, but also metaphor may play a role in conversion – for example, in verbs converted from nouns denoting animals (e.g. *to parrot*), which include a metaphorical mapping between the domain of animals and the domain of humans.

Mititelu et al. (2021) use a sample of 17 thousand deverbal nouns and denominal verbs, including conversion pairs, from WordNet. The annotation of semantic relations is taken from a stand-off file (Fellbaum et al. 2009). The results show that conversion is pervasive across all semantic categories, but not equally distributed. The hierarchical structure of WordNet also makes it possible to enrich the semantic classification with information about which general classes of verbs (e.g. verbs of communication, verbs of motion, verbs of change) and nouns (e.g. artifact, attribute, person) are utilized most frequently in each category and how they cluster together.

Huyghe (2021) presents a database of French deverbal nouns, which contains four thousand verb/noun pairs extracted from a large web corpus of French. Nouns created by suffixation and conversion are included along with their semantic description, which presents yet another type of semantic classification for deverbal nouns. The semantic labels include two types of information: the noun's ontological type and relational type. The former specifies the type of entity that the noun denotes (e.g. ARTIFACT, EVENT, PROPERTY, STATE). The latter specifies the relation of the noun's meaning to the motivating verb's meaning, based on semantic roles of the verb's arguments (e.g. AGENT, CAUSE, DESTINATION, INSTRUMENT). Many of the nouns are polysemous and therefore classified into multiple categories.

3 Method

3.1 Semantic categories used in the annotation experiment

The set of semantic categories used in the present experiment has been compiled on the basis of Bauer et al.'s (2013) and Cetnarowska's (1993) inventories, but also the other approaches reviewed above have been taken into consideration; yet without any ambition to include all the categories or to capture all the nuances discussed. While the previous approaches either focus exclusively on one of the directions, or they postulate different categories for the deverbal and denominal conversions, our classification puts semantic relations that are closely related under one label. This means that the same label is applied to the noun that has been converted from a verb as well as to the noun that is the source for the verbs if the nouns show a similar semantic relation to the respective verbs based on the paraphrases. For instance, Bauer et al.'s category RESULT, which is listed among the meanings of deverbal nouns (with the paraphrase "the

outcome of V-*ing*"; e.g. *divorce* = "the result of divorcing"), and their denominal category RESULTATIVE ("to make into X" as in *to bundle* = "to make into bundles") were put under the category RESULT in the present analysis.

Ten semantic categories that are used to capture the semantic relationship within the conversion pairs in the annotation are listed in alphabetical order in Table 1. Each category is followed by paraphrases for the deverbal and/or denominal formations, and by an example conversion pair.

Table 1: Semantic categories used in the annotation experiment

Semantic category (meaning paraphrase)	Example conversion pair		Corresponding category in Bauer et al. (2013)
ACTION (action of V-*ing*/to perform N)	*útok* 'attack'	*útočit* 'to attack'	PERFORMATIVE (denominal) EVENT, INSTANCE (deverbal)
ADDED (something added through V-*ing*/to provide with N)	*olej* 'oil'	*olejovat* 'to oil'	ORNATIVE (denominal)
AGENT (someone who performs V-*ing*/to act like N or to be N)	*král* 'king'	*kralovat* 'to reign'	STATIVE, SIMILATIVE (denominal) AGENT (deverbal)
ANIMAL (–/to be similar to N)	*křeček* 'hamster'	*křečkovat* 'to hoard up'	–
INSTRUMENT (something used for V-*ing*/ to use N)	*hrábě* 'rake'	*hrabat* 'to rake'	INSTRUMENT (denominal) INSTRUMENTAL (deverbal)
PLACE (a place to which something is V-*ed*/to place something to N)	*láhev* 'bottle'	*lahvovat* 'to bottle'	LOCATIVE (denominal)
REMOVED (something removed through V-*ing*/to remove N)	*skalp* 'scalp'	*skalpovat* 'to scalp'	PRIVATIVE (denominal)
RESULT (result of V-*ing*/to make into N)	*vrstva* 'layer'	*vrstvit* 'to layer'	RESULTATIVE (denominal) RESULT, PRODUCT (deverbal)
STATE (the state of being V-*ed*/to bring someone into the state of N)	*šok* 'shock'	*šokovat* 'to shock'	STATE (denominal)
TIME (the time spent V-*ing*/–)	*noc* 'night'	*nocovat* 'to overnight'	–

Remarks on the semantic categories listed in the table:

i. Bauer et al.'s categories SIMILATIVE ("to act like X", *to chauffeur*) and STATIVE ("to be X", *to hostess*) are not distinguished because the boundary between them appears fuzzy with regards to the pairs under analysis; cf. also Efthymiou (2013) on Greek.

ii. There does not seem to be any deverbal category in Bauer et al.'s classification that would accommodate such deverbal pairs in which the objects or qualities are named after the action by which they are ADDED or REMOVED (in contrast, the Czech data contain such examples).

iii.   While both denominal and deverbal cases of the STATE category are found in the Czech data, a corresponding category is found only among denominal (not among deverbal) categories in Bauer et al. ("the state of V-*ing* or being V-*ed*", e.g. *fear*). Deverbal instances from English are listed by Cetnarowska.

iv.   The TIME category does not have a counterpart in Bauer et al.'s classification, but corresponds to DURATION verbs in Clark and Clark's typology of denominal verbs. Cetnarowska lists several deverbal examples from English (e.g. *fall* and *freeze* in the category THE PERIOD OF V-*ing*).

Although our classification does not contain a residual category (such as MISCELLANEOUS verbs in Clark and Clark's study), it does not mean that all conversion pairs fit equally well into the proposed categories; see the analysis of the results of our annotation experiment in §4.

3.2 Compilation of the data

Conversion pairs to which the semantic categories are to be assigned were extracted from a 100-million corpus of written Czech (*SYN2015*, Křen et al. 2015). The study is limited to conversion pairs which are made up of a suffixless noun and a corresponding verb that differs from the noun solely by the verbal inflections.[2] Suffixless nouns both without inflectional endings (cf. (1) and (2) above) and with the endings -*a*, -*e/-ě* and -*o* (in (3) to (5)) are included. In all examples the nouns and verbs are prefixless and share only the root, except for (3) where the common part also includes a prefix.

(3)   *za-bav-i-t*   >   *zá-bav-a*
       PREF-enjoy-PFV-INF   PREF-enjoy-NOM.SG.F
       'to entertain'   'entertainment'

(4)   a.   *lyž-ova-t*   <   *lyž-e*
             ski-IPFV-INF   ski-NOM.SG.F
             'to ski'   'ski'
       b.   *kup-ova-t*   >   *koup-ě*
             buy-IPFV-INF   buy-NOM.SG.F
             'to purchase'   'purchase'

(5)   *zlat-i-t*   <   *zlat-o*
       gold-IPFV-INF   gold-NOM.SG.N
       'to gild'   'gold'

The conversion pairs were searched in the *SYN2015* corpus using the *Morfio* tool (Cvrček & Vondřička 2013), which searches the corpus data for pairs or n-tuples of lexemes matching provided regular expressions. A query was formulated for each of the Czech verbal theme types in relation to nouns without endings and with the respective endings. In the example queries in Table 2, the regular expression `{.+[:Kons:]}` represents the string shared by the noun and the verb (i.e. one or more characters followed by a consonant), which in the first query is

---

[2] Nouns containing non-native elements that have not been reanalyzed as suffixes in Czech were considered to be suffixless in the present study (cf. -*ment* in *dokument* 'document' or -*or* in *monitor* 'monitor'), even if these strings may function as suffixes in the source languages.

followed by no other character in the noun and by four characters in the verb, while in the second query it is followed by *-o* in the noun and by two characters in the verb. Example pairs that match these regular expressions are provided, including those (in grey) that are to be excluded from the sample. The tool allows for vowel and/or consonant alternations within the string shared by the noun and the verb.

Table 2: Example queries to search for conversion pairs with the *Morfio* tool.
Discarded examples in shaded font

noun	verb	noun/verb pairs matching the queries		
`{.+[:Kons:]}`	`{.+[:Kons:]}....`	*boj*	–	*bojovat*
		'fight'		'to fight'
		odhad	–	*odhadnout*
		'estimate'		'to estimate.PFV'
		odhad	–	*odhadovat*
		'estimate'		'to estimate.IPFV'
		ředitel	*–*	*ředitelovat*
		'director'		'to be a director'
		růž	*–*	*růžovět*
		'blush'		'to turn pink'
`{.+[:Kons:]}o`	`{.+[:Kons:]}..`	*kouzlo*	–	*kouzlit*
		'magic'		'to do magic'
		zlato	–	*zlatit*
		'gold'		'to gild'
		vápno	–	*vápnit*
		'quicklime'		'to lime'
		dělo	*–*	*dělit*
		'cannon'		'to divide'

To ensure that the sample actually contains only conversion pairs with the suffixless nouns and the directly related verbs connected by synchronously available semantic relations, the extraction by the *Morfio* tool was followed by a manual check to remove:

i.   nouns and verbs that have a common etymological origin but are not linked by a semantic relation perceived in synchrony (e.g. *příběh* 'story' vs. *přiběhnout* 'to come running'),
ii.   pairs that matched a particular query only formally without a semantic link between the noun and verb (cf. *dělo* 'cannon' – *dělit* 'to divide' in Table 2),
iii. pairs in which the noun contains a suffix, which became a part of the verb (cf. *ředitel* 'director' – *ředitelovat* 'to be a director' in Table 2),
iv.   pairs in which the suffixless noun and the corresponding verb were parts of the same morphological family but were not directly related (cf. *růž* 'blush' and the verb *růžovět* 'to turn pink' in Table 2; the suffix *-ov-* in the verb reveals that it is motivated by the adjective *růžový* 'pink').

If a particular noun was listed with two verbs that contain different thematic suffixes conveying different grammatical aspects (cf. the entries with the noun *odhad* 'estimate' in Table 2), these two entries were merged into one (i.e. *odhad* 'estimate' – *odhadnout* : *odhadovat* 'to estimate').

Moreover, in this step the existence of an aspectual counterpart with a different theme was checked for all verbs in the data and, if it was found, the counterpart was added to the data. These steps resulted in a dataset of more than 2,200 entries, which either consisted of a suffixless noun and a directly corresponding verb, or of a suffixless noun and a pair of verbs with different themes. By using the Pythonic `random.sample()` built-in function, 800 out of the 2,200 entries were randomly selected for the parallel semantic analysis.

The semantic relationship between the noun and the verb was assessed on the basis of the concordances documented in the *SYN2015* corpus. For each conversion pair, a random sample was extracted that contained 50 sentences with the noun and 50 sentences with the verb (or verbs).[3] If one or both of the members of the conversion pair were documented with less than 50 hits in the corpus, all available concordances were included in the sample for annotation.

3.3 Parallel annotation of the semantic relations within the conversion pairs

The semantic annotation was carried out by two annotators in parallel on the same set of concordances. At the time of the annotation, both annotators were Master students of English and Czech philology; one of them with extensive experience with linguistic analysis of conversion data for English and Czech, while the second annotator without previous experience with word-formation data.

In the default case, all sentences with a particular noun documented the same reading that was related to one particular reading of the corresponding verb. For instance, the noun *olej* 'oil' in all the analyzed sentences referred to a substance added to a thing by the action expressed by the verb *olejovat* 'to oil'. The annotators were instructed to assign the semantic category in the sentences containing the item that was assumed to be motivated based on the aspect-changing strategy, in order to keep the annotation as simple as possible and for an easy evaluation. That is, for conversion pairs where the denominal direction is assumed, the sample of sentences containing the verbs was annotated, while for conversion pairs where the deverbal direction is assumed, annotation was carried out on sentences containing the noun. With the example conversion pair, the ADDED category was assigned to the verb *olejovat* 'to oil', which is assumed to be motivated by the noun because it is not capable of changing the grammatical aspect by replacing the theme.

If the sentences documented more than a single sense–sense relation, more than one label was used. For instance, the noun *řez* 'cut' was annotated as an ACTION in some concordances while in others as a RESULT of the activity expressed by the verbs *říznout* : *řezat* 'to cut'. In a similar vein, in the conversion pair *slina* 'saliva' – *slinit* 'to salivate/to wet something with saliva', the verb is understood as an activity of producing what is referred to by the noun (RESULT) or of adding what is referred to by the noun (ADDED), depending on the context of the particular sentence.

Those readings that obviously were not input or output of conversion with respect to the other pair member were marked as irrelevant; cf. the noun *plat* 'pay' that relates to the verb *platit* only in the meaning of 'to pay', but not as 'to be valid'. Some sentence samples did not contain any concordance in which the noun and the verb were semantically related to each other – after discarding these, parallel annotation was obtained for 739 pairs.

---

[3] If the noun corresponds to two verbs with different themes, sentences containing both verbs were included in the sample.

The parallel annotation as well as the final disambiguated dataset, which is described in the next section, have been made publicly available as part of the repository item http://hdl.handle.net/11234/1-5142 in the LINDAT/CLARIAH-CZ repository under the CC BY-NC-SA 4.0 license.

4 Results of the annotation experiment

4.1 Evaluation of the parallel annotation

The semantic relationship between the noun and the verb in 739 conversion pairs was determined by each annotator in 30,277 concordances. In total, the annotators agreed on semantic categories assigned to 22,461 out of 30,277 sentences, i.e. on 74% of the sentences under analysis.

The inter-annotator agreement was also calculated by using Krippendorff's (1970, 2012) alpha coefficient, which reduces the effect of chance agreement when evaluating an analysis by more than two annotators and/or an analysis in which the annotators use more than two labels.[4] Krippendorff's alpha reaches values 1 and lower: 1 indicates perfect agreement and thus perfect reliability of the annotation; values between 1 and 0.800 indicate excellent agreement; values lower than 0.800 but higher than or equal to 0.667 signal fair agreement; values between 0.667 and 0 indicate low agreement; and 0 or values below 0 indicate complete absence of reliability or disagreement going beyond coincidence. When applied to the results of the present annotation experiment, Krippendorff's alpha coefficient reaches the value 0.79, which speaks of a fair agreement between the two annotators, bordering on excellent.

The individual semantic labels are not equally distributed, i.e. some are assigned more frequently than others. The semantic category used most frequently by both annotators is ACTION, followed by RESULT, INSTRUMENT, and ADDED. The only difference concerns the AGENT and PLACE categories – while AGENT is applied more frequently than PLACE by Annotator A, Annotator B preferred PLACE over AGENT. The labels TIME, REMOVED, and ANIMAL were the least frequent categories in both annotations.

Table 3 shows a confusional matrix of the choices of the two annotators. The semantic categories assigned by Annotator A are listed in columns, while the categories chosen by Annotator B are in rows. The numbers of sentences whose classification the annotators agreed upon are displayed in the darkest gray in the cells on the diagonal. The other cells contain the numbers of cases on whose classification the annotators disagreed. A scale from dark gray to white is used to visualize how large the set of disagreeing cases is relative to the cases on which the annotators agreed in a given category.

There is a perfect match between the annotators in assigning the labels TIME and ANIMAL. However, these two categories are marginal, which makes the agreement on their annotation less informative. On the other hand, there are noticeable disagreements in the annotation of the following pairs (Annotator A vs. Annotator B, respectively in what follows):

---

[4] In our study, Krippendorff's alpha was preferred over the more common Cohen's kappa coefficient (Cohen 1960) because of the number of categories used in the parallel annotation. While Cohen's kappa and the related metrics BAK and PABAK (bias-adjusted kappa and prevalence-adjusted bias-adjusted kappa, respectively; Byrt et al. 1993) work well for measuring inter-annotator agreement in binary classification tasks, in the experiment presented in this paper the task was to agree on one of ten categories, which is better evaluated by the more robust Krippendorff's measure.

ACTION vs. RESULT, ACTION vs. STATE, RESULT vs. ADDED, RESULT vs. INSTRUMENT, RESULT vs. STATE, and RESULT vs. PLACE; the sources of disagreement are analyzed in the following section. Absolute disagreement is in white in the table (all instances classified as REMOVED by Annotator B were assessed as RESULT by Annotator A). Also white are cells filled with an 'x' which indicates that the particular combination did not occur in the parallel annotation. The total number of sentences to which Annotator A and Annotator B assigned a given category is shown for each category in the bottom row and in the first column, respectively.

Table 3: Semantic categories assigned by Annotator A (columns) vs. by Annotator B (rows)

Annotator B	ACTION (12379)	RESULT (7637)	INSTRUMENT (3137)	STATE (2153)	ADDED (2208)	AGENT (1263)	PLACE (1396)	TIME (100)	ANIMAL (1)	REMOVED (3)
ACTION (10958)	9908	529	162	190	108	30	31	x	x	x
RESULT (7984)	1462	5417	221	298	363	55	168	x	x	x
INSTRUMENT (4371)	162	657	2512	53	461	92	432	x	x	2
STATE (2828)	765	426	6	1563	17	1	50	x	x	x
ADDED (1961)	14	403	236	49	1258	x	x	x	x	1
AGENT (1093)	58	x	x	x	x	1035	x	x	x	x
PLACE (882)	10	106	x	x	1	50	715	x	x	x
TIME (100)	x	x	x	x	x	x	x	100	x	x
ANIMAL (1)	x	x	x	x	x	x	x	x	1	x
REMOVED (99)	x	99	x	x	x	x	x	x	x	x

Annotator A

4.2 Sources of disagreement

The analysis of the instances where the annotators differed pointed to three sources of disagreement, namely to the coarse-grained nature of the semantic categories, to different levels of granularity in the annotators' decisions, and to the lack of clarity in one of the annotation guidelines.

The fact that, given the potentially unlimited semantic range of conversion, the semantic categories are relatively few and the boundaries between them are blurred relates to most of the differences observed in the parallel annotation. The delimitation of individual categories in the continuous semantic space is done using paraphrases. However, as has been previously noted in the literature, several equally good paraphrases may be possible for a single word (even used in a given context), especially when it comes to more abstract meanings; cf. Gottfurcht's (2008: 16) discussion of the verb *satirize*, which can be paraphrased as "perform a satire" (PERFORMATIVE), "convert something into a satire" (RESULTATIVE), "apply satire to some situation" (ORNATIVE), or "use a satire" (INSTRUMENTAL). On the other hand, there may be instances where none of the paraphrases seems to be applicable, also reported in literature (and usually classified into some kind of miscellaneous category). For example, Valera (2020: 325) reports that 7.89% of the total number of concordances that were analysed could not be classified into previously defined categories.

The fuzzy boundaries between some of the categories can be demonstrated by a closer look into the pairs of categories with the relatively largest degree of disagreement in our annotation experiment:

i. ACTION VS. RESULT is mostly found in polysemous words which allow both meanings, e.g. *výhra* 'win, victory' – *vyhrát* : *vyhrávat* 'to win'. The concordance is often insufficient to distinguish which of the meanings is used, in fact, many concordances are ambiguous and allow for both readings. In such cases, the decision for one or the other category is to a certain extent arbitrary, and it may even be more appropriate to use a joint ACTION–RESULT category.

ii. ACTION VS. STATE appears in pairs in which the verb is neither a prototypical action, nor a prototypical state, e.g. *zátěž* 'burden' – *zatížit* : *zatěžovat* 'to burden'. States and actions are points on a single spectrum, and although they can be differentiated by certain criteria (e.g. the degree of control and intention of the agent or the level of dynamicity), there will always be borderline cases.

iii. ADDED VS. RESULT can be found in pairs where the activity denoted by the verb affects things that already exist, e.g. *náklad* 'load, cargo' – *naložit* : *nakládat* 'to load'. In the final annotation, in which all disagreements were resolved (see §4.3), the dividing line is drawn so that cases where a new detail, characteristics, etc. is added to the original item without its original form being fundamentally altered are treated as ADDED, whereas cases where the input is changed into something new are classified as RESULT.

iv. ADDED VS. INSTRUMENT often occurs in pairs like *maska* 'mask' – *maskovat* 'to mask, put on a mask', where adding something means using it. This occurs especially in pairs which include nouns referring to concrete things. This type of ambiguity is discussed already in Clark & Clark (1979: 778): "Take *Ned leashed the dog*: is *leash* a locatum verb ('Ned caused the dog to have a leash on it'), or an instrument verb ('Ned caused the dog to be restrained by doing the act one would normally expect to do to a dog with a leash')?". The authors offer several criteria for distinguishing the two classes, but admit that such sentences can be ambiguous or vague.

v. RESULT VS. STATE appears in pairs where the noun denotes a state that is caused by the action/process expressed by the verb, e.g. *návyk* 'custom, habit' – *navyknout* : *navykat* 'to get accustomed'. In other words, the noun is a resulting state and both categories are justified. Huyghe's (2021) classification would allow both to be included in the label (STATE as an ontological type, RESULT as a relational type), however, in the classification used here, these are defined as two separate categories.

vi. INSTRUMENT VS. PLACE occurs most frequently in pairs containing verbs of motion, such as *vstup* 'entrance' – *vstoupit* : *vstupovat* 'to enter', where the noun can be paraphrased both as "what is used for entering" and "the place through which one enters". The set of semantic categories which is used does not include categories typical for the semantic roles of verbs of movement, such as PATH, SOURCE, DESTINATION (in contrast to e.g. Huyghe 2021, who includes these categories among the relational types), and therefore the annotators had to decide which of the available labels is the closest. One annotator consistently used the INSTRUMENT label in such cases, while the second one used the PLACE label. In the final annotation, those locations that serve for the performance of the activity are labeled as INSTRUMENTs.

vii. RESULT VS. PLACE is also connected to the absence of categories specific for the verbs of movement (e.g. *směr* 'direction' – *směřovat* 'to direct'), but also to a set of other, quite diverse verbs with an unclear preferred paraphrase, such as *konec* 'end/ending' ("the place where something ends/the result of something ending") – *končit* 'to end' or *blog* 'blog' ("the place where one blogs/the result of someone blogging") – *blogovat* 'to blog'.

viii.   REMOVED vs. RESULT is essentially limited to two verbs of bodily secretion (*potit se* 'to sweat' – *pot* 'sweat', *močit* 'to urinate' – *moč* 'urine'), which one annotator classified based on the paraphrase "to remove sweat, urine from the body", but the other annotator classified them based on the paraphrase "to produce sweat, urine". Due to this difference, the REMOVED category has zero agreement.

Another source of divergent analyses – closely related to the previous one – was that in some cases the annotators distinguished the meanings of the words in the concordances with a different level of granularity. There are several conversion pairs in which one annotator chose a single category for all sentences in the respective sample, while the other distinguished two or more categories in different concordances (e.g. *pomoc* 'help' – *pomoci* : *pomáhat* 'to help' as ACTION vs. ACTION+AGENT; *výstava* 'exhibition' – *vystavit* : *vystavovat* 'to exhibit' as ACTION vs. ACTION+PLACE). The issue of what should be considered a discrete meaning and what should only be considered a contextual modification of a single meaning is central to research on polysemy and has been raised in connection with conversion, for example, by Cetnarowska (1993: 96). It seems that the strategy of distinguishing categories based on meaning paraphrases is not fully sufficient to resolve this issue.

The third source of disagreement was the instruction to specify semantic categories for the item that is considered motivated based on the aspect strategy of the verb in the conversion pair. In accordance with this instruction, the semantic categories in the pair *cesta* 'path/trip' – *cestovat* 'to travel' were determined with the verb, but one annotator chose the INSTRUMENT category in relation to the reading of the noun as a 'path' (as "something used for traveling"), the other chose the category ACTION in relation to the meaning of a 'trip' (as "the action of traveling"). Both readings were documented with the noun in the corpus sample, but the method did not allow it to be captured.

4.3 Analysis of disambiguated data

In the next step, all annotator disagreements were systematically resolved in favour of one of the labels to create a final version of the data. In addition to deciding on individual instances, the ANIMAL category was cancelled and the respective instances were included in the AGENT group. The reason for this revision was that although we originally expected ANIMAL nouns to be in a more complicated relation to verbs than nouns and verbs in the AGENT category, it has turned out during the annotation that the pairs do not basically differ from AGENT verbs (cf. *plazit (se)* 'to creep' = "to act, in particular, to move like a reptile"). Our final treatment of ANIMAL nouns as AGENT nouns is in line, for instance, with Clark & Clark' (1979: 773–774) decision of handling verbs with "animal parents" as AGENT verbs, but in contrast to Martsa's (2013: 155–167) delimitation of ANIMAL verbs as a separate group.

The final disambiguated dataset includes 739 conversion pairs annotated in a total of 30,277 concordances. On average, 41 concordances were assigned a semantic label for each conversion pair. Less than five annotated concordances are available for 49 conversion pairs, but most conversion pairs (500) have at least 45 annotated concordances available. The distribution of individual semantic categories in the final data is quantified in Table 4.

Table 4: The number of concordances with the individual semantic labels in the disambiguated dataset

Semantic label	Absolute occurrence	Relative occurrence
ACTION	12,192	40.3%
RESULT	7,603	25.1%
INSTRUMENT	3,689	12.2%
STATE	2,297	7.6%
ADDED	2,266	7.4%
AGENT	1,133	3.8%
PLACE	895	3.0%
REMOVED	102	0.3%
TIME	100	0.3%
Total	30,277	100.0%

The disambiguated dataset also makes it possible to assess the number of meanings identified with individual conversion pairs and to analyze the patterns of polysemy. In 80% of the conversion pairs, all concordances were assigned the same semantic category. For the other conversion pairs, either two semantic categories (18% of the pairs) or three semantic categories (2% of the conversion pairs) were required to annotate all the sampled concordances. Table 5 provides the absolute counts of conversion pairs that were assigned a particular single semantic category or a particular combination of multiple semantic categories. With a total of 906 semantic labels assigned to the 739 conversion pairs, the final dataset documents an average polysemy rate of 1.23 meanings per conversion pair.

If combinations of semantic labels annotated with individual conversion pairs are counted to assess the polysemy of the pairs, one can simply see that most of the conversion pairs were annotated by using a single label, out of which ACTION was the most frequent. However, there were also relatively frequent combinations of multiple labels, some of which were also often the source of inter-annotator disagreement (ACTION+RESULT, INSTRUMENT+RESULT, ADDED+RESULT, PLACE+RESULT). It seems that these categories not only cause uncertainty in categorizing a certain meaning of a verb/noun in context, but also coexist as separate meanings of one polysemous noun/verb. This supports a hypothesis that in these cases, polysemy may be a result of semantic shift, which often happens between closely related meanings that may be ambiguous in a given context.

The patterns of polysemy which appear multiple times in the data are often those which include the ACTION category combined with one or two other categories. In these cases, the ACTION category was often found in the majority of the sentences annotated for a given conversion pair (this is especially true for ACTION+INSTRUMENT, ACTION+AGENT; INSTRUMENT and AGENT are often marginal, in that they appear in only a few sentences). In the context of semantic shift, frequent combinations of the actional meaning with another, non-actional meaning seem to support the traditional Czech view (e.g. Daneš et al. 1967: 261; Štícha 2018: 440) that the actional meaning is primary for deverbal nouns, and concrete meanings have evolved from the actional meaning through the process of semantic change. A similar claim has been made about English nouns converted from verbs by Cetnarowska (1993: 118; cf.

semantic formation in Štekauer 1996: 99). However, our analysis cannot support this claim generally, because there are many conversion pairs which only exhibit non-actional readings without also having an actional reading.

The two categories which appear together significantly more frequently than any other pair of categories, both in the patterns of polysemy and in inter-annotator disagreements, are ACTION and RESULT. A similar closeness was observed for these categories also in affixation, for example in a recent study of French deverbal nouns (Salvadori & Huyghe 2023), where the authors examined which meanings are often expressed by the same affix and therefore hypothesized to be semantically related. Although the semantic classification is different than the one used in this study, the semantic categories of EVENT-TRANSPOSITION and ARTIFACT-RESULT, which have been found among the most frequently co-occurring categories, are to a large extent comparable to our ACTION and RESULT categories. This provides some evidence for the conceptual relatedness of the two categories as a general phenomenon, not necessarily restricted to a specific word-formation process in a given language.

Table 5: Semantic labels and their combinations in the final version of the data

Semantic label(s)	# of pairs	Example	
ACTION	202	*boj* 'fight'	– *bojovat* 'to fight'
RESULT	151	*segment* 'segment'	– *segmentovat* 'to segment'
INSTRUMENT	78	*zvon* 'bell'	– *zvonit* 'to ring the bell'
ADDED	64	*zlato* 'gold'	– *zlatit* 'to gild'
ACTION+RESULT	57	*zisk* 'gain/profit'	– *získat* : *získávat* 'to gain/to profit'
STATE	45	*nuda* 'boredom'	– *nudit* 'to bore'
ACTION+INSTRUMENT	27	*sonda* 'probe'	– *sondovat* 'to probe'
AGENT	24	*král* 'king'	– *kralovat* 'to reign'
PLACE	20	*garáž* 'garage'	– *garážovat* 'to garage'
ACTION+AGENT	9	*návštěva* 'visit'	– *navštívit* : *navštěvovat* 'to visit'
INSTRUMENT+RESULT	9	*náplň* 'filling/contents'	– *naplnit* : *naplňovat* 'to fill'
ACTION+INSTRUMENT+RESULT	7	*odkaz* 'reference/legacy'	– *odkázat* : *odkazovat* 'to refer/to leave'
ADDED+RESULT	6	*mlha* 'mist'	– *mlžit* 'to add mist/to produce mist'
PLACE+RESULT	6	*řada* 'group/line'	– *řadit* 'to classify/to line up'

ACTION+ADDED	4	*podpis* 'signature'	–	*podepsat* : *podepisovat* 'to sign'
ACTION+PLACE	3	*úkryt* 'hiding/shelter'	–	*ukrýt* : *ukrývat* 'to hide/to shelter'
ACTION+STATE	3	*úleva* 'relief'	–	*ulevit* : *ulevovat* 'to relieve'
ADDED+INSTRUMENT	3	*obklad* 'compress/tiling'	–	*obložit* : *obkládat* 'to cover'
REMOVED	2	*pot* 'sweat'	–	*potit* 'to sweat'
TIME	2	*noc* 'night'	–	*nocovat* 'to overnight'
ADDED+STATE	2	*záruka* 'guarantee'	–	*zaručit* : *zaručovat* 'to guarantee'
AGENT+RESULT	2	*perla* 'pearl'	–	*perlit* 'to sparkle'
ACTION+PLACE+RESULT	2	*západ* 'sunset/west'	–	*zapadnout* : *zapadat* 'to set/to fall'
AGENT+INSTRUMENT	1	*posila* 'support'	–	*posílit* : *posilovat* 'to support'
INSTRUMENT+REMOVED	1	*led* 'ice'	–	*ledovat* 'to ice/to remove ice'
REMOVED+RESULT	1	*prach* 'dust'	–	*prášit* 'to dust/to produce dust'
RESULT+STATE	1	*zápal* 'enthusiasm/inflammation'	–	*zapálit* : *zapalovat* 'to ignite'
ACTION+ADDED+RESULT	1	*nátěr* 'paint'	–	*natřít* : *natírat* 'to paint'
ACTION+AGENT+INSTRUMENT	1	*náhrada* 'substitute/substitution'	–	*nahradit* : *nahrazovat* 'to substitute'
ACTION+AGENT+RESULT	1	*výprava* 'expedition/setting'	–	*vypravit* : *vypravovat* 'to set (out)'
ACTION+INSTRUMENT+PLACE	1	*východ* 'rise/exit/east'	–	*vyjít* : *vycházet* 'to rise/to go out'
ACTION+INSTRUMENT+STATE	1	*nájem* 'rent/rental'	–	*najmout* : *najímat* 'to rent'
ADDED+PLACE+RESULT	1	*síť* 'net/network'	–	*síťovat* 'to equip with networks/to put into a net/to create a network'
AGENT+RESULT+STATE	1	*otrava* 'annoying person/ poisoning/boredom'	–	*otrávit* : *otravovat* 'to annoy/to poison/to bore'

5 Conclusions

The annotation experiment presented in this study was carried out with two main aims. Firstly, to apply semantic classification (based mostly on Bauer et al. 2013 and Cetnarowska 1993) to a large sample of Czech conversion data and to test whether it is possible to achieve good agreement between two independent annotators; secondly, to see whether the frequent sources of annotation disagreements, as well as the final results of the annotation after all disagreements were resolved, can reveal something about the relationships that the semantic categories have to one another.

As for the first aim, fair inter-annotator agreement (Krippendorff's alpha coefficient = 0.79) was reached and the annotators agreed in the majority of cases in all semantic categories (except the marginal REMOVED category). Disagreements were found to be connected mostly to the number of categories which, after a closer look into the data, were found to have fuzzy boundaries between one another.

As for the second aim, we take the fuzziness of boundaries manifested in inter-annotator disagreements as well as the frequent co-occurrence of certain semantic categories in one conversion pair (polysemy patterns) as evidence in support of the conceptual closeness of the semantic categories. Some of the patterns of polysemy were combinations of categories which also appeared as sources of inter-annotator disagreement, and combinations of the actional meaning with another, non-actional meaning are relatively frequent. This speaks in support of polysemy being a result of semantic shift between conceptually close categories in these cases.

Acknowledgements

References

Adams, Valerie. 1973. *An Introduction to Modern English Word-Formation*. London: Longman. https://doi.org/10.4324/9781315504254

Bauer, Laurie. 1983. *English Word-formation*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139165846

Bauer, Laurie & Lieber, Rochelle, & Plag, Ingo. 2013. *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198747062.001.0001

Bednaříková, Božena. 2009. *Slovo a jeho konverze*. Olomouc: Univerzita Palackého.

Bozděchová, Ivana. 2016. Czech. In Müller, Peter O. & Ohnheiser, Ingeborg & Olsen, Susan & Rainer, Franz (eds.), *Word-Formation: An International Handbook of the Languages of Europe*, Vol. 4, 2872–2891. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110379082-025

Byrt, Ted & Bishop, Janet & Carlin, John B. 1993. Bias, prevalence and kappa. *Journal of Clinical Epidemiology* 46(5). 423–429. https://doi.org/10.1016/0895-4356(93)90018-V

Cetnarowska, Bożena. 1993. *The Syntax, Semantics and Derivation of Bare Nominalisations in English*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Cetnarowska, Bożena. 1996. Constraints on suffixless derivation in Polish and English: The case of action nouns. In Kardela, Henryk & Szymanek, Bogdan (eds.), *A Festschrift for Edmund Gussmann from his friends and colleagues*, 15–28. Lublin: The University Press of the Catholic University of Lublin.

Clark, Eve V. & Clark, Herbert H. 1979. When nouns surface as verbs. *Language 55*(4). 767–811. https://doi.org/10.2307/412745

Cohen, Jacob. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 20(1). 37–46.

Cvrček, Václav & Vondřička, Pavel. 2013. Nástroj pro slovotvornou analýzu jazykového korpusu. In *Grammar and corpora/Gramatika a korpus 2012*, 1–10. Hradec Králové: Gaudeamus.

Daneš, František & Dokulil, Miloš & Kuchař, Jaroslav. 1967. *Tvoření slov v češtině 2: Odvozování podstatných jmen*. Prague: Nakladatelství ČSAV.

Dokulil, Miloš. 1962. *Tvoření slov v češtině 1: Teorie odvozování slov*. Praha: Academia.

Dokulil, Miloš. 1968a. Zur Frage der Konversion und verwandter Wortbildungsvorgänge und -beziehungen. *Travaux Linguistiques de Prague* 3. 215–239.

Dokulil, Miloš. 1968b. Zur Frage der sog. Nullableitung. In Brekle, Herbert E. & Lipka, Leonhard (eds.), *Wortbildung, Syntax und Morphologie: Festschrift zum 60. Geburtstag von Hans Marchand am 1. Oktober 1967*, 55–64. The Hague: De Gruyter Mouton. https://doi.org/10.1515/9783111349152-008

Dokulil, Miloš. 1982. K otázce slovnědruhových převodů a přechodů. *Slovo a slovesnost* 43. 257–271.

Dokulil, Miloš & Horálek, Karel & Hůrková, Jiřina & Knappová, Miloslava & Petr, Jan. 1986. *Mluvnice češtiny 1: Fonetika, fonologie, morfonologie a morfematika, tvoření slov*. Praha: Academia.

Efthymiou, Angeliki. 2013. How Many Factors Influence the Meaning of Denominal and Deadjectival Verbs? In ten Hacken, Pius & Thomas, Claire (eds.), *The Semantics of Word Formation and Lexicalization*, 225–246. Edinburgh: Edinburgh University Press. https://doi.org/10.3366/edinburgh/9780748689606.003.0012

Fellbaum, Christine & Osherson, Anne & Clark, Peter E. 2009. Putting Semantics into WordNet's "Morphosemantic" Links. In Vetulani, Zygmunt & Uszkoreit, Hans (eds.), *Human Language Technology: Challenges of the Information Society*, 350–358. Berlin: Springer. https://doi.org/10.1007/978-3-642-04235-5_30

Gottfurcht, Carolyn A. 2008. *Denominal Verb Formation in English*. Evanston, Illinois: Northwestern University. (Doctoral dissertation.)

Hledíková, Hana. 2022. *Conversion in English and Czech: A corpus study of semantic relations between nouns and verbs*. Prague: Charles University. (Master thesis.) http://hdl.handle.net/20.500.11956/174483

Huyghe, Richard. 2021. Building a lexical database to investigate the semantics of French deverbal nouns. In Namer, Fiammetta & Hathout, Nabil & Lignon, Stéphanie & Ševčíková, Magda & Žabokrtský, Zdeněk (eds.), *Proceedings of the Third International Workshop on Resources and Tools for Derivational Morphology (DeriMo 2021)*, 21–29.

Krippendorff, Klaus. 1970. Estimating the reliability, systematic error, and random error of interval data. *Educational and Psychological Measurement* 30 (1). 61–70. https://doi.org/10.1177/001316447003000010

Krippendorff, Klaus. 2012. *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks: Sage.

Křen, Michal & Cvrček, Václav & Čapka, Tomáš & Čermáková, Anna & Hnátková, Milena & Chlumská, Lucie & Jelínek, Tomáš & Kováříková, Dominika & Petkevič, Vladimír & Procházka, Pavel & Skoumalová, Hana & Škrabal, Michal & Truneček, Petr & Vondřička, Pavel & Zasina, Adrian J. 2015. *SYN2015: A Representative Corpus of Written Czech*. Prague: Charles University. (http://www.korpus.cz) (Accessed 2022-03-10)

Manova, Stela. 2011. *Understanding Morphological Rules: With Special Emphasis on Conversion and Subtraction in Bulgarian, Russian and Serbo-Croatian*. Dordrecht: Springer. https://doi.org/10.1007/978-90-481-9547-3

Marchand, Hans, 1969. *The categories and types of present-day English word-formation: A synchronic-diachronic approach*. Munich: C. H. Beck.

Martsa, Sándor. 2013. *Conversion in English: A Cognitive Semantic Approach*. Cambridge: Cambridge Scholars.

Mititelu, Verginica & Leseva, Svetlozara & Stoyanova, Ivelina. 2021. Semantic Analysis of Verb-Noun Derivation in Princeton WordNet. *Proceedings of the 11th Global Wordnet Conference*. 108–117.

Plag, Ingo. 1999. *Morphological productivity: Structural constraints in English derivation*. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110802863

Plag, Ingo. 2003. *Word-formation in English*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511841323

Salvadori, Justine & Huyghe, Richard. 2023. Affix polyfunctionality in French deverbal nominalizations. *Morphology* 33(1). 1–39. https://doi.org/10.1007/s11525-022-09401-4

Ševčíková, Magda. 2021a. Action nouns vs. nouns as bases for denominal verbs in Czech: A case study on directionality in derivation. *Word Structure* 14. 97–128. https://doi.org/10.3366/word.2021.0181

Ševčíková, Magda. 2021b. Bezpříponová substantiva a vyjadřování vidového protikladu u příbuzných sloves. *Slovo a slovesnost* 82. 263–288.

Štekauer, Pavol. 1996. *A theory of conversion in English*. Frankfurt am Main: Peter Lang.

Štekauer, Pavol & Valera, Salvador & Körtvélyessy, Lívia. 2012. *Word-Formation in the World's Languages*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511895005

Štícha, František. 2018. *Velká akademická gramatika spisovné češtiny 1: Morfologie: Druhy slov/Tvoření slov*. Prague: Academia.

Valera, Salvador. 2015. Conversion. In Müller, Peter O. & Ohnheiser, Ingeborg & Olsen, Susan & Rainer, Franz (eds.), *Word-Formation: An International Handbook of the Languages of Europe*, Vol. 1, 322–339. Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110246254-019

Valera, Salvador. 2020. Semantic Patterns in Noun-to-Verb Conversion in English. In Körtvélyessy, Lívia & Štekauer, Pavol (eds.), *Complex Words,* 311–334. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108780643.017

Valera, Salvador. 2023. The semantics of noun-to-verb zero-derivation in English and Spanish. *Zeitschrift für Sprachwissenschaft* 42(1). 153–180. https://doi.org/10.1515/zfs-2022-2016

Magda Ševčíková
Charles University
Faculty of Mathematics and Physics
Institute of Formal and Applied Linguistics
Malostranské nám. 25
118 00 Praha, Czech Republic
[*sevcikova@ufal.mff.cuni.cz*]

Hana Hledíková
Charles University
Faculty of Mathematics and Physics
Institute of Formal and Applied Linguistics
Malostranské nám. 25
118 00 Praha, Czech Republic
[*hana.hledikova@ufal.mff.cuni.cz*]

Lukáš Kyjánek
Charles University
Faculty of Mathematics and Physics
Institute of Formal and Applied Linguistics
Malostranské nám. 25
118 00 Praha, Czech Republic
[*kyjanek@ufal.mff.cuni.cz*]

Anna Staňková
Charles University
Faculty of Arts
Institute of Czech Language and Theory of Communication
nám. Jana Palacha 2
116 38 Praha, Czech Republic
[*anna.stankova@ff.cuni.cz*]