

Inflectional and derivational paradigm of verbs in Czech: the role of the category of aspect

Magda Ševčíková, Jarmila Panevová, Patrice Pognan

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

ParadigMo2017
20. 6. 2017





Contents

- 1 Introduction
- 2 Aspect in the verb-to-verb derivation
 - Aspect as an inflectional category
 - Verb-to-verb derivation in Czech
- 3 Aspect in DeriNet database data
 - DeriNet: basic facts
 - Aspect in organizing the database data
- 4 Future work and conclusions



Intro: Inflection and derivation of verbs in Czech

- inflection strictly separated from derivation in grammar books of Czech
 - inflectional morphemes cumulate grammatical meanings
 - derivation by agglutinative morphemes
- inflection of verbs
 - categories of person, number, tense, mood, verbal voice
plus gender
plus aspect

kreslila '(she) drew.impf / was drawing'

kresl-i-l-a

draw-THEME-PST.IND.ACT-3SG.FEM

nakreslila '(she) drew.pf'

na-kresl-i-l-a

PF-draw-THEME-PST.IND.ACT-3SG.FEM



Introduction: Aspect of verbs in Czech

- verbal aspect in Slavic languages
 - “grammatical aspect”
 - Comrie (1976), Meřčuk (1994, 2016); Kopečný (1962), Komárek et al. (1986), Komárek (2006)
- in Czech
 - imperfective (impf) vs. perfective (pf) verbs
 - a (lexico-)grammatical category in the theoretical description of Czech
 - the aspectual counterpart is a form of the particular verb
 - expressed by derivational affixes
- in the paper
 - not to challenge the inflectional status of the category of aspect
 - to present aspect as an important feature to be used when describing derivation of Czech verbs



An example: inflection and derivation of *nakreslit_{pf}* to draw

- inflectional forms (200+ in total):
 - ind. pres./pret. act.: *nakreslím, nakreslíš, nakreslila jsem, nakreslili, ...*
 - cond. pres./past act.: *nakreslili bychom, byly by nakreslily, ...*
 - transgr. pres./past act.: *nakreslíc, nakreslivši, ...*
 - ind. pres./pret. pass.: *je nakreslen, byla nakreslena, ...*
 - cond. pres./past act.: *byli bychom nakresleni, bylo by bývalo nakresleno, ...*
- lexemes derived from *nakreslit_{pf}* to draw:
 - *nakreslit* → *nakreslený_A* drawn
 - *nakreslit* → *nakreslivší_A* having drawn
 - *nakreslit* → *nákres_N* draft
 - *nákres_N* draft → *nákresna_N* projection plane



An example: inflection and derivation of *nakreslit_{pf}* to draw

- inflectional forms (200+ in total):
 - ind. pres./pret. act.: *nakreslím, nakreslíš, nakreslila jsem, nakreslili, ...*
 - cond. pres./past act.: *nakreslili bychom, byly by nakreslily, ...*
 - transgr. pres./past act.: *nakreslíc, nakreslivši, ...*
 - ind. pres./pret. pass.: *je nakreslen, byla nakreslena, ...*
 - cond. pres./past act.: *byli bychom nakresleni, bylo by bývalo nakresleno, ...*
- lexemes derived from *nakreslit_{pf}* to draw:
 - *nakreslit* → *nakreslený_A* drawn
 - *nakreslit* → *nakreslivši_A* having drawn
 - *nakreslit* → *nákres_N* draft
 - *nákres_N* draft → *nákresna_N* projection plane



The category of aspect: the meaning and form

- impf vs. pf verbs
 - continuous vs. complex representation of the same action
- aspect expressed by derivational affixes
 - the pf verb and impf verb differ in the suffix:
 - *dá-t_{pf}* to give – *dá-va-t_{impf}* to give
 - *skoč-i-t_{pf}* to jump – *ská-k-a-t_{impf}* to jump
 - *koup-i-t_{pf}* to buy – *kup-ova-t_{impf}* to buy
 - *štěk-nou-t_{pf}* to bark – *štěk-a-t_{impf}* to bark
 - perfectivization of impf verbs by a prefix:
 - *kreslit_{impf}* to draw – *na-kreslit_{pf}* to draw
 - *kreslit_{impf}* to draw – *pře-kreslit_{pf}* to copy
- pure aspectual pair of verbs
 - an impf and a pf verb that share the same lexical meaning
 - pure perfectivizing prefix vs. other perfectivizing prefixes



The category of aspect: the meaning and form

- impf vs. pf verbs
 - continuous vs. complex representation of the same action
- aspect expressed by derivational affixes
 - the pf verb and impf verb differ in the suffix:
 - *dá-t_{pf}* to give – *dá-va-t_{impf}* to give
 - *skoč-i-t_{pf}* to jump – *ská-k-a-t_{impf}* to jump
 - *koup-i-t_{pf}* to buy – *kup-ova-t_{impf}* to buy
 - *štěk-nou-t_{pf}* to bark – *štěk-a-t_{impf}* to bark
 - perfectivization of impf verbs by a prefix:
 - *kreslit_{impf}* to draw – *na-kreslit_{pf}* to draw
 - *kreslit_{impf}* to draw – *pře-kreslit_{pf}* to copy
- pure aspectual pair of verbs
 - an impf and a pf verb that share the same lexical meaning
 - pure perfectivizing prefix vs. other perfectivizing prefixes



The pure perfectivizing prefix

- often one of many prefixes that combine with the particular verb:
kreslit_{pf} to draw – *na-kreslit_{pf}* to draw
vs. *do-kreslit_{pf}* to finish drawing, *pře-kreslit* to redraw, *ob-kreslit* to copy, *při-kreslit* to add by drawing, ...
- formally different with different verbs:
kreslit_{impf} – *na-kreslit_{pf}* to draw
dělat_{impf} – *u-dělat_{pf}* to do
chválit_{impf} – *po-chválit_{pf}* to praise
- pf verb with a pure perfectivizing prefix cannot form a secondary imperfective:
nakreslit_{pf} – **nakreslovat* vs. *překreslit_{pf}* – *překreslovat_{impf}* to copy
BUT *napsat_{pf}* – ?*napisovat* (1 hit in SYN corpus), *pochválit_{pf}* – *pochvalovat_{impf}*



Verb-to-verb derivation in Czech

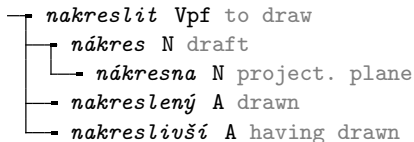
- expression of aspect just one of the functions of affixation of verbs
- prefixation
 - perfectivization of impf verbs: *kreslit_{impf}* – *na-kreslit_{pf}* to draw
 - modification of the lexical meaning of an impf or pf verb:
kreslit_{impf} to draw – *pře-kreslit_{pf}* to copy
skočit_{pf} to jump – *vy-skočit_{pf}* to jump out
- suffixation
 - imperfectivization of pf verbs (or vice versa): *skočit_{pf}* – *skákat_{impf}* to jump
 - deriving iteratives from impf verbs: *skákat_{impf}* – *skákávat_{impf.iter}* to jump
 - deriving secondary impf from prefixed pf verbs: *překreslit_{pf}* – *překreslovat_{impf}* to copy



Lexical database DeriNet [<http://ufal.mff.cuni.cz/derinet>]

- lexical resource specialized in derivation of Czech
- lexemes extracted from the infl. dictionary (Hajič – Hlaváčová 2013)
 - 1,012k lexemes in total
- derived words connected with the particular base word
 - semi-automatic methods
- pairs of derivationally related lexemes form a derivational tree

*nakreslit*_{Vpf} to draw → *nákres*_N draft
*nákres*_N → *nákresna*_N projection plane
*nakreslit*_{Vpf} → *nakreslený*_A drawn
*nakreslit*_{Vpf} → *nakreslivší*_A having drawn



Word Formation Latin (Litta et al. 2016), Démonette (Hathout – Namer 2014), DerivBase.Hr (Šnajder et al. 2014), language-independent approach (Baranes – Sagot 2014), DerivBase (Zeller et al. 2013), CELEX (Baayen et al. 1995), ...



Identifying and organizing derivationally related verbs

- task: to interlink derivationally related verbs in DeriNet data
 - to delimit verbs with the common root
 - to specify criteria for a coherent and linguistically plausible organization
- typically, tens of verbs with a common root
 - complicated relations among the verbs
- aspectual counterparts NOT excluded from the derivational tree due to many open questions:
 - which verbs form an aspectual pair
 - which of the aspectual counterparts is the marked member
 - what to do with the “missing links” created in this way

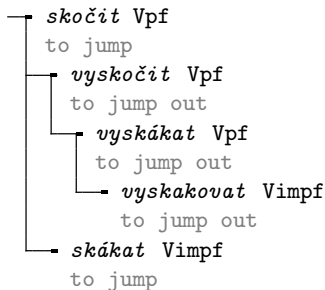
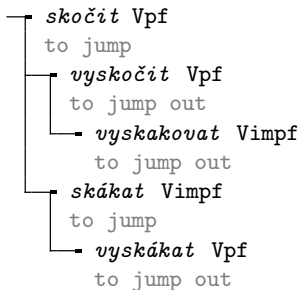
{skoč / skák / skak} “to jump”

skočit_{pf}, *skákat_{impf}*, *vyskočit_{pf}*, *přeskočit_{pf}*, *naskočit_{pf}*, *vyskákat_{pf}*,
přeskákat_{pf}, *přeskakovat_{impf}*, *vyskakovat_{impf}*, *naskakovat_{impf}*, ...



Creating the trees

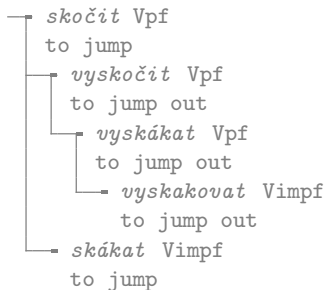
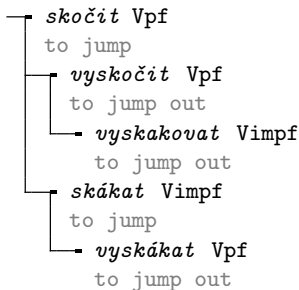
- the inner organization is not implicitly included in the data
 - several “correct” solutions





Creating the trees

- the inner organization is not implicitly included in the data
 - several “correct” solutions





Designing derivational trees: adopted criteria

- the unmarked member of the aspectual pair is the root of the tree
 - in aspectual pairs derived by suffixation: pf or impf verb
skočit_{pf} → *skákat_{impf}* to jump, *štěkat_{impf}* → *štěknout_{pf}* to bark
 - in aspectual pairs derived by prefixation: impf verb
psát_{impf} → *napsat_{pf}* to write
- prefixed pf verbs are derived from non-prefixed counterparts (pf or impf):
skočit_{pf} → *naskočit_{pf}* / *vyskočit_{pf}* / *poskočit_{pf}* / ...
skákat_{impf} → *naskákat_{pf}* / *vyskákat_{pf}* / *poskákat_{pf}* / ...
- secondary impf verbs are derived from prefixed pf verbs:
naskočit_{pf} → *naskakovat_{impf}* to hop on
- iterative verbs are derived from respective impf verbs:
skákat_{impf} → *skákávat_{impf.iter}* to jump,
zaskakovat_{impf} → *zaskakovávat_{impf.iter}* to substitute



Implementation: the verb-to-verb derivation in DeriNet data

- 1 aspectual pairs of verbs extracted from an existing valency dictionary of Czech (Lopatková et al. 2015)
 - suffix string of the related verbs identified
- 2 a list of app. 50 prefixes used in the verb-to-verb derivation was compiled
- 3 valency data searched for prefixed verbs related to the aspectual pairs
> 3,100 verbs found, organized into 660 derivational trees
- 4 further verbs searched for in DeriNet data, added into the trees
> 23k new derivational relations between 27k verbs identified



Results: statistics of DeriNet 1.4

- 23k new derivational relations between 27k verbs identified
 - V-to-V derivation: 21k in DeriNet 1.3 > 44k in DeriNet 1.4

lexemes in total	1,012k	deriv. links in total	772k
nouns (N)	442k	V to A	203k
adjectives (A)	354k	A to N	162k
adverbs (D)	161k	A to D	157k
verbs (V)	55k	N to A	106k
		V to N	58k
		V to V	44k
		N to N	40k
		A to A	2k



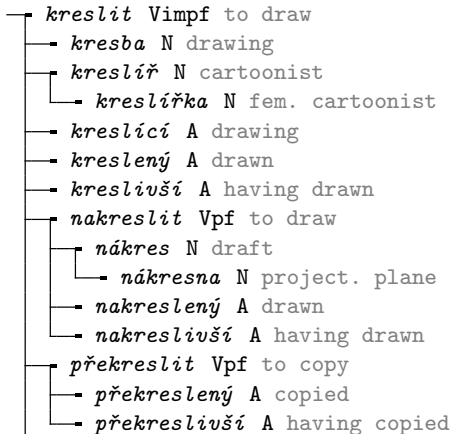
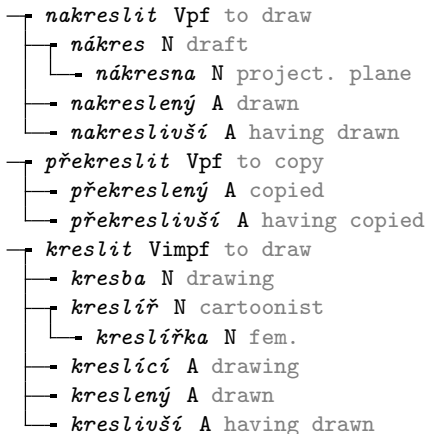
Results: statistics of DeriNet 1.4

- 23k new derivational relations between 27k verbs identified
 - V-to-V derivation: 21k in DeriNet 1.3 > **44k** in DeriNet 1.4

<u>lexemes in total</u>	<u>1,012k</u>	<u>deriv. links in total</u>	<u>772k</u>
nouns (N)	442k	V to A	203k
adjectives (A)	354k	A to N	162k
adverbs (D)	161k	A to D	157k
verbs (V)	55k	N to A	106k
		V to N	58k
		V to V	44k
		N to N	40k
		A to A	2k

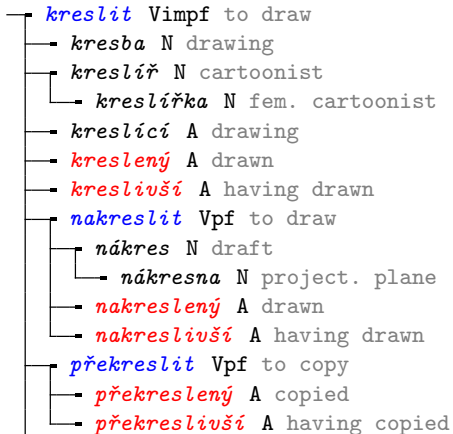
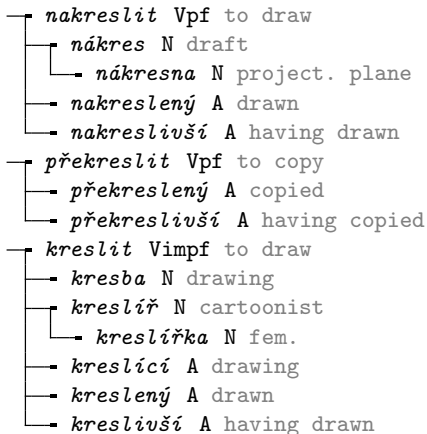


Results: changes in derivational trees of verbs





Results: changes in derivational trees of verbs





Future work, conclusions

- large derivational trees searched for repeating patterns (subtrees)
 - > a data-base research into derivational paradigms (Štekauer 2014)?
- derivation of verbs can be analyzed with respect to the tree structure
 - the root verb has a richer set of derivatives than derived verbs
 - derivation of derived verbs is limited to regular derivatives (esp. adjectives in Czech) based on inflected forms
 - derivation vs. lexical meaning and syntactic features
 - derived verbs form a syntactic construction instead of forming an own derivative
*kreslit – kreslíř vs. nakreslit – *nakreslíř – kreslíř nákresu*



Future work, conclusions (cont.)

- aspect as an inflectional category
 - inflectional category of Czech verbs
 - expressed by derivational morphemes
- in DeriNet data
 - formal features preferred: aspectual pairs of verbs as derivationally related
 - aspect to an important feature in modelling verb-to-verb derivation in Czech
- arguing in favor of an integrating approach to inflection and derivation of Czech verbs

- Baayen, R. H. et al. (1995): *The CELEX lexical database* (release 2). Data/software. Philadelphia, PA: LDC.
- Baranes, M. – Sagot, B.: A Language-Independent Approach to Extracting Derivational Relations from an Inflectional Lexicon. *LREC* 2014:2793–2799.
- Comrie, B. (1976): *Aspect*. Cambridge: Cambridge University Press.
- Hajič, J. – Hlaváčová, J. (2013): *MorfFlex CZ*. LINDAT/CLARIN digital library at Institute of Formal and Applied Linguistics, Charles University in Prague, <http://hdl.handle.net/11858/00-097C-0000-0015-A780-9>
- Hathout, N. – Namer, F.: Démonette, a French Derivational Morpho-Semantic network. *LiLT* 2014:11, 125–168.
- Komárek, M. (2006): *Příspěvky k české morfologii*. Olomouc: Periplum.
- Komárek, M. a kol. (1986): *Mluvnice češtiny 2. Tvarosloví*. Praha: Academia.
- Litta, E. et al. (2016): *Formatio formosa est. Building a Word Formation Based Lexicon for Latin*. In: *CLiC-it 2016*: 185–189.
- Lopatková, M. a kol. (2017): *Valenční slovník českých sloves Vallex*. Praha: Karolinum.
- Mel'čuk, I. A. (1994): *Cours de morphologie générale*. Vol. 2. Montréal: Presses de l'Université de Montréal.
- Mel'čuk, I. A. (2016): Formaľnoje vyraženiye vida v ruskom jazyke. *Russkij jazyk v naučnom osvješčenii* 31, pp. 9–23.
- Šnajder, J. et al.: DerivBase.Hr: A High-Coverage Derivational Morphology Resource for Croatian. *LREC* 2014: 3371–3377.
- Štekauer, P. (2014): Derivational paradigms. In: R. Lieber – P. Štekauer (eds.): *The Oxford Handbook of Derivational Morphology*. Oxford: Oxford University Press, pp. 354–369.
- Zeller, B. et al.: DERivBase: Inducing and evaluating a derivational morphology resource for German. *ACL* 2013: 1201–1211.

<http://ufal.mff.cuni.cz/derinet>