



# The Unbearable Lightness of Morph Classification

Vojtěch John and Zdeněk Žabokrtský<sup>(✉)</sup>

Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics,  
Charles University, Malostranské náměstí 25, 118 00 Prague, Czech Republic  
zabokrtsky@ufal.mff.cuni.cz

**Abstract.** In light of the recent push for the creation and unification of large morphologically annotated resources, there is a call for (preferably language-independent, low-resource) methods of morph classification. This paper reports on a pilot experiment on morph classification of the Czech language. We have performed two experiments - root morph recognition and complete morph classification. By exploiting simple quantitative methods and - in some cases - available Czech morphological resources, we have achieved morph-level precision of respectively 96.7% and 88.3%.

**Keywords:** morphology · morphological analysis · morph classification · DeriNet

## 1 Introduction

The standard NLP task of morphological segmentation, i.e. dividing words into sequences of the smallest possible meaning-bearing units called morphemes<sup>1</sup>, has recently seen a fair share of renewed interest. Recent development in morphological analysis of underresourced languages and/or multilingual morphological analysis (e.g. [1] or [9]). The state-of-the-art methods of morphological segmentation are however usually based on neural networks [1], and hence are neither easily generalizable to languages with fewer resources nor straightforwardly usable for further morphological analysis. As there is a push for the creation and unification of morphologically annotated resources [3, 26], and as adding morphological and syntactic information seems to improve the quality of machine translation for morphologically rich and underresourced languages [8], a need arises for new methods for language-independent low-resource methods of morph classification.

Up to now, there has been relatively little attention paid to the task. Classification of morphs aims to classify the individual morphs given already segmented words, the possible granularity of the classification ranging from the simple binary distinction *free morpheme - bounded morpheme* to e.g. the very complex and fine-grained Leipzig glossing rules [7]. To our best knowledge, there has been no recent attempt to tackle automated classification of morphs of the Czech language apart from [5], where the

<sup>1</sup> Or morphs, if we speak about particular forms that appear in words.

DeriNet derivational lexicon is used for root morph identification; however, this is just quickly mentioned in the article and the results are neither evaluated nor discussed. Furthermore, to this date, there is a very limited amount of reasonably-high-quality morphologically annotated Czech data available (apart from the UniMorph project [3], whose unsuitability for our purposes we will discuss in the next section).

In this paper, we report on a pilot experiment concerned with the classification of morphs of the Czech language. Firstly, we try to show how the root and non-root morphs of pre-segmented words are surprisingly well distinguishable using simple quantitative approaches and a small corpus of other pre-segmented words; we also propose classification methods that exploit existing Czech derivational resources. Secondly, we report on a work in progress in which we propose to use the root morph identification methods as a basis for complete morpheme classification of pre-segmented words, exploiting the available Czech derivational and morphological resources.

It will be noticed that the proposed methods differ both in their generality and the demands on resources. Namely, we also use two additional resources: the Czech derivational network DeriNet [24, 25] and MorfFlex [12] (for lemmatization and part-of-speech tagging). Nevertheless as according to [5] similarly constructed derivational networks are available for at least 11 languages and lemmatization and part-of-speech tagging are among the best-explored topics in NLP, we feel confident that these additional requirements are not so stringent as to make possible adaptations or generalizations of our approach to other languages too costly.

## 2 Related Work

### 2.1 Terminology

In the following text, we use the terminology as described in [26]. A morpheme is the smallest (in the sense of non-subdivisible) sequence of graphemes associated with a definite meaning. In individual words, morphemes are present in particular forms – morphs [13]. The morphs can be further characterized. We can distinguish free morphemes that can be used as separate words), and bound morphemes that cannot<sup>2</sup>. Alternatively, the morphemes are either lexical (with more or less general lexical meaning) or grammatical (with inflectional meaning).

Based on these two distinctions, we distinguish root morphemes (free lexical morphemes; e.g. *kůň*, *plav*), derivational affixes (bound lexical morphemes, e.g. *pro-*, *-tel*), inflectional affixes (bound grammatical morphemes, e.g. *nej-*, *-ý*) and function words (free grammatical morphemes, e.g. *s*, *a*). According to their position relative to the root morpheme (resp. morph), we may divide affixes into prefixes preceding the root morph, interfixes between the root morphs or suffixes following the root. We could also take into account postfixes - morphemes that appear after inflectional suffixes.

<sup>2</sup> It should be noted that these terms, while universally used, are not completely suitable for the Czech language, where many of the morphs cannot be used as separate words because words that contain them demand an inflectional affix in every form (e.g. “pín”, the root morph for “odepínat” – “to bind”, which alone is not a valid word in Czech); even though they would otherwise be classified as root morphs.

## 2.2 Data Resources

There are several kinds of relevant data resources. First, there are morphological lexicons. For Czech, there are two main lexicons, unfortunately not available in machine tractable form: *Retrográdní morfematický slovník češtiny* [21] and *Bázový morfematický slovník češtiny* [20]. Furthermore, the Czech language is included in the UniMorph project [3]. Nevertheless, the Czech data in UniMorph, automatically extracted from MorfFlex [12], are segmented in a way that is incompatible with our proposed morph classification. Namely, the root in the UniMorph segmentation actually seems to be the lemma or the root of the derivational tree (often including non-inflectional prefixes). For the morphological segmentation, there are 38 000 manually segmented Czech words in the data used for the SIGMORPHON 2022 shared task [1].

In addition, several of the available derivational resources already contain at least some kind of morphological segmentation and classification. A survey of this type of resource can be found in [26]. A good example of the kind of morphological resource we have in mind is the manually created CroDeriV lexicon [22]. It contains over 14 000 Croatian words (except two nouns all verbs) segmented to morphs. The morphs are classified as prefixes, stems, suffixes or endings.

The granularity of the classification included in the data differs: the Dictionary of Morphemes of the Russian Language [16] contains over 74 000 segmented Russian lemmas with labeled root morphs (and not affixes). Furthermore, for many languages, either no such resources exist or - as they are created automatically or semi-automatically - they are not straightforwardly usable as a source of gold data, such as in the case of the German morphological derivational lexicon Derivbase [27] or the recent multilingual derivational and inflectional database MorphyNet [2]. This database also includes Czech data; they are however incompletely and quite often inaccurately segmented. Therefore, however useful this resource might prove to be for practical purposes, we are reluctant to employ it for our pilot experiment (especially as the gold data).

The Czech derivational network DeriNet [25] contains a rough morphological segmentation (resp. for 250 000 of the lemmas the root morphs are labeled). The methods by which the root morphs are labeled are nevertheless much similar to some of the methods we try (and therefore cannot be used as the gold data); furthermore, the relevant articles [5, 25] do not mention the final accuracy of the method.

## 2.3 Morphological Segmentation and Classification

Some of the classical approaches to morphological analysis either already include, or could be straightforwardly extended to include, the classification of morphs. Thus Goldsmith's unsupervised morphological segmentation [10] uses minimum description length and several simple heuristics to generate candidate stems and suffixes. Unsupervised morphology induction like Schone and Jurafsky [19] or more recently the word-embedding-based induction proposed by Soricut and Och [23] use automatically extracted affixes for morphological rules induction.

For languages for which well-annotated and sufficiently large resources of the abovementioned type are available, supervised machine learning can be used for both the tasks - morphological segmentation and classification of morphs. Recently, Bolshakova and Sapin [6] employed a neural model for morphological segmentation and classification of Russian, achieving over 90% word-level classification accuracy.

In comparison, the state-of-the-art results for the Czech language are not so promising. The Czech derivational network DeriNet has been used for morphological segmentation (and partial classification) [5] with the achieved word-level segmentation accuracy of 58.9 % (the number is only illustrative - the accuracy of root morph recognition was not measured). This might be caused by the lack of available relevant Czech data (almost no available data for morpheme classification, till recently ([1, 26]) also for segmentation).

### 3 Data

We use four data resources in total. First, we use a small set of fully manually segmented and annotated words (316 words in total), which we further subdivide into the dev set and test set (each containing 158 words).<sup>3</sup> The morphs are annotated by their type, similarly to the CroDeriV [22]. The classes are as follows:

- **R** - root morphs,
- **P** - derivational prefixes,
- **S** - derivational suffixes,
- **I** - inflectional affixes,
- **N** - interfixes,
- **O** - postfixes.

Secondly, mainly as “training data” (for feature extraction of morphs), we use 10 438 manually segmented Czech non-compound words with manually selected root morphs. Thirdly, in some of our experiments, we use the DeriNet Czech derivational lexicon. It contains over 1M Czech lexemes connected by over 800 000 derivational relations. We also use the MorfFlex lexicon [12], which contains 125 348 899 simple lemma - tag - form triples. The tags, as described in [11], are very fine-grained and contain morphological as well as syntactic information.

### 4 Evaluation

There are two possible levels of evaluation—word-level evaluation and morph-level evaluation. The word-level evaluation measures are less fine-grained, but they offer some desirable properties (e.g. the instances might be weighted by the number of occurrences of a given word in the corpus, or unweighted - giving each word equal weight. Imagine two extreme scenarios - A) half of the annotations is completely right, the other half is completely wrong; and B) half of the morphs in every word is annotated right and the other half is wrong. While e.g. unweighted word-level precision of the first example would be 50% and the precision of the second one would be 0%, while the morph-level precision of both examples would be 50%.

For both of our experiments, we have selected five simple evaluation measures: Word-level accuracy and morph-level precision, recall, F-measure and accuracy. It

<sup>3</sup> [https://github.com/johnvojtech/morph\\_analysis/blob/main/data/completed\\_hand\\_annotation.tsv](https://github.com/johnvojtech/morph_analysis/blob/main/data/completed_hand_annotation.tsv).

should be noted however that since we use a very small test set, the word-level accuracy, while included for illustration and completeness, should be nevertheless regarded with some caution.

## 5 Experiment 1: Root Morph Selection

### 5.1 Methods

In our first experiment, our goal is to identify the root morph of the word. We started with three baseline heuristics. We also - as a 0-th baseline - tried to label every morph as an affix (since affixes are more common than roots) - *r0*. For all following heuristics, we compute the morph features on test set + training data, i.e. the 10k segmented words, without the annotation of the root morph.

First, we selected all the longest morphs of the segmented words. Secondly, we selected the morph with the least occurrences in our dictionary on the intuition that, since compounds is infrequent in Czech, root morphs are usually combined with (a limited number of) affixes, while affixes can combine with (a big number of) root morphs. This method was used in the annotation of DeriNet [25] with the first one as the tiebreaker. Thirdly, we estimate the left and right conditional entropy of the morphs and select the morph with the smallest difference between the two. This was motivated by the observation that while the roots usually appear in between two affixes at the beginning/end of the word, the affixes usually appear only on one side of the root morph, while on the other side, there is either the beginning/end of the word or another affix; thus, one would expect the difference between the left and right entropy to be quite high in the case of affixes and quite low in the case of root morphs. Finally, we combine all these three heuristics - we normalize them (so all three sum to 1) and minimize their unweighted sum; instead of the length we use  $1/\text{length}$  (so that we may minimize it).

Further, we experimented with methods based on the derivational network DeriNet [25]. First, for each word, we found the unmotivated lemma (or “root lemma”) of the derivational tree<sup>4</sup> (in DeriNet) and all its children, computed the edit distance between each morph and these words and selected the morph with the shortest edit distance. This we use either by itself (*r5*) or in combination with the previous three heuristics (*r6*) in the same way as in (*r4*). Lastly, instead of taking into account only the unmotivated lemma of the current word’s derivation tree and its children, we computed the longest common substring of all descendants of the unmotivated lemma, (including replacing any character with a wildcard to - very roughly - deal with possible allomorphy; thus e.g. the common substring of “sit” and “sat” would be “s?t”) and then apply *r6*.

In all of the (*r5*–*r7*), if the processed word is not found in the DeriNet, we use the *r4* method. For comparison with a supervised approach, we also trained the CRF tagger implemented in NLTK [4] on the training data (with annotated root morphs); that is, we treat the segmented words as sentences and the morphs as tagged words.

---

<sup>4</sup> A tree structure, where derivationally related words are organised according to derivational history; the root of the tree represents the unmotivated lemma.

## 5.2 Results

As we announced in the introduction, the simple quantitative methods gave surprisingly good results (see Table 1). Every method apart from just taking the longest morphs (r1) achieved higher precision than the CRF tagger. As all of the methods (also apart from r1) were restricted to selecting exactly one root morph, the achieved F-measure is also surprisingly good (94.4%). We can also note that even the best of the fully-unsupervised methods (i.e. not using DeriNet) achieves comparable F-measure with the supervised CRF tagger.

**Table 1.** Evaluation of root morph identification.

	Word-level accuracy	Morph-level precision	Morph-level recall	Morph-level f-measure	Morph-level accuracy
CRF tagger	89.9	90.4	<b>94.7</b>	92.5	96.3
r0	–	–	–	–	68.9
r1	68.4	81.1	84.5	82.7	89.5
r2	86.1	92.4	86.9	89.6	93.9
r3	86.1	93.9	86.9	89.6	93.9
r4	88.6	94.9	89.3	92.0	95.4
r5	88.0	94.3	88.7	91.4	95.0
r6	90.5	96.8	91.1	93.9	96.4
r7	<b>91.1</b>	<b>97.4</b>	91.7	<b>94.4</b>	<b>96.8</b>

## 6 Experiment 2: Morph Classification

### 6.1 Methods

In our second experiment, we try to expand our root morph recognition methods to a fully-fledged morphological classifier. The most important part of this is the distinction between root morphs, the derivational affixes and the inflectional affixes.

### 6.2 Baselines

We have implemented two baseline morph classifiers, supervised and unsupervised one. As our first, supervised baseline (Baseline 1), we assign to each morph the tag that is most commonly associated with the morph in our development set; if it is not present there, we label it as a root morph (it is the most frequent label in the dev set). As our second, unsupervised baseline (Baseline 2), we designed two versions of a simple unsupervised heuristics-based classifier. First, we decide for each morph in a given word whether it could be a derivational affix, inflectional affix or a root morph (by heuristics described in the following subsections). In the second iteration, we assign to each of the positions the tag that is most common for the morph in the processed data. In the B version we consider as possible inflectional morphs only the first and last morphs of the word.<sup>5</sup>

<sup>5</sup> In all our baseline solutions we ignore the interfixes and postfixes.

**Derivational Affixes.** Using DeriNet, we take the morphs in which the segmented word fully differs from the root lemma; i.e., we compute the minimum edit distance between the word and the root lemma, and label the characters that would be added or rewritten during the minimal edit. We regard all the morphs, that consist only of such characters, as derivational morphs. This heuristics is very rough in that most of the Czech lemmas contain an inflectional affix, which in this way could also get classified as a derivational affix.

**Inflectional Affixes.** For each tag in MorfFlex [12], we take a thousand word-forms corresponding to the same tag; also for every word form present in the data, we take all the other word forms. The inflectional affixes are those that are (more or less) common for common tags, but different for the different forms. Namely, we try to extract an ending common for all the words tagged by the same tag; if that fails, we consider every ending common for at least one-fifth of the examples. If even that fails, we try to find the longest uncommon substring (including wildcard characters) for all the forms of the segmented word.

**Roots.** For root morph recognition we use the DeriNet-based *r7* method from the previous section.

### 6.3 Finetuning CRF Taggers

To exploit the transition probabilities between words, we have (as in the first experiment) used the bidirectional LSTM-CRF tagger as described in [14]. The CRF tagger used in NLTK does not permit finetuning, so we have used the implementation from the *bi-lstm-crf* Python package [15]. First, we have trained it on the 10k training set with the manually annotated root morphs. Then, we finetuned it on the small development set (which contains 158 annotated words). This approach (Semi-supervised CRF), however, presupposes the rather large data with annotated root morphs. However, since we have developed methods of root morph recognition, we could use them for creating the training data.

In our second CRF tagger-based experiment (Supervised CRF), we have used the large training data stripped of the manual annotation and annotated automatically, using our most successful DeriNet-based method (*r7*). Thirdly, we have trained the CRF tagger only on the dev set (Small CRF). Finally, we have pre-trained the CRF tagger on the training data annotated by both of our baseline solutions (and again finetuned on the dev set).

### 6.4 Evaluation and Results

In the second experiment, we evaluate only the word-level and morph-level accuracy (Table 2). The results were somewhat surprising. First of all, no version of the CRF tagger was better than one of our unsupervised baselines, which has achieved 88 % morph-level accuracy. Secondly, while both pretraining the CRF tagger on root identification and on data annotated by the baseline methods seems to have significant impact

**Table 2.** Evaluation of full classification of morphs.

	Morph-level accuracy	Word-level accuracy
Baseline 1	83.7%	55%
Baseline 2A	76.6%	18%
Baseline 2B	<b>88.3%</b>	<b>70%</b>
Small CRF	72.9%	42 %
Supervised CRF	75.7%	44%
Semi-supervised CRF	79.7%	54%
Baseline 1 + CRF	84.6%	63%
Baseline 2A + CRF	61.2%	37%
Baseline 2B + CRF	80.8%	57%

**Table 3.** Example output (Baseline 2B)

Segmented word	Assigned signature	Correct signature
styd nou t	RSI	RSI
při prav i t	PRSI	PRSI
pol ovič n í	RSSI	RSSI
ne roz trh a l i	IPRSSI	IPRSSI
po prsk a t	PRSI	PRSI
maž	R	R
z kypř e t	PRSI	PRSI
o hod n ot i t	PRSSSI	PRSSSI
obe še l	PRS	PRI
sáh l a	SSI	RSI

on their accuracy, there is no clear correspondence between overall quality of the pre-training data and the overall quality of the CRF tagger results. Thirdly, there is a big difference in accuracy between the two versions of our unsupervised baseline.

Sometimes, the taggers make mistakes that could be fairly easily filtered out (but always not so easily corrected), as having a suffix before a root, a word without a root (see the last example in Table 3) or a sequence like “Root - Prefix - Suffix”. Introduction of simple rules might therefore significantly increase the final accuracy. One such example causes the large difference between the accuracy of the two versions of Baseline 2. Closer look at the data reveals that most of the errors of the version A of the unsupervised baseline solution consisted in misidentification of derivational suffixes as inflectional, which might be easily filtered out by the restriction on the position of the inflectional affix, as used in the B version; the real accuracy of the A version might however be higher, as in some cases the identification of derivational suffixes in the manually annotated test set is spurious (e.g. *u klid n i l i* is assigned the signature *PRSSSI* in the test set and *PRSI* by the baseline solution; but the suffix *-l*, expressing past tense, might be said to be inflectional).



**Table 4.** Errors in morph classification (except rare categories, i.e. infixes and postfixes)

System/Error (Correct:Assigned)	R:S	R:P	R:I	S:R	S:P	S:I	I:R	I:P	I:S	P:R	P:I	P:S
Baseline 1	0	0	1	17	2	11	6	0	<b>28</b>	9	0	2
Baseline 2A	0	3	1	5	3	<b>74</b>	8	11	14	4	1	1
Baseline 2B	0	4	0	5	5	4	8	11	<b>18</b>	4	0	1
Small-data CRF	11	<b>43</b>	5	19	0	22	11	5	19	8	0	1
Supervised CRF	<b>36</b>	16	0	24	3	18	1	6	10	9	0	4
Semi-supervised CRF	7	<b>25</b>	0	<b>25</b>	2	17	6	6	6	4	0	0
Baseline 1 + CRF	6	9	6	<b>17</b>	4	11	6	5	4	7	0	6
Baseline 2A + CRF	7	<b>42</b>	7	<b>42</b>	12	39	8	11	19	9	0	2
Baseline 2B + CRF	3	14	8	16	0	<b>19</b>	6	5	0	7	1	1

## 7 Conclusion

We have shown that applying simple quantitative methods on comparatively small and/or unannotated segmented data is sufficient for a high-quality root morph identification in Czech and that these results can be further improved by exploiting the DeriNet derivational lexicon. In our second experiment, we used our root morph identification methods to create training data and to train an LSTM-CRF tagger. It appears that the quality of the output can be increased by pre-training the tagger on root morph identification or morph data classified by a good-enough baseline solution. Furthermore, the simple supervised baseline was as good as the CRF taggers, while one of the unsupervised baselines has been significantly more accurate.

In the future, we would like to better utilize Czech resources like MorfFlex and DeriNet either for further morphological analysis (e.g. the derivational affixes would appear in many derivational trees in DeriNet but on the lower levels, while root morphs would appear in only a few trees and on all levels; most the inflectional affixes would not appear at all or at no specific level). Also, the morphological tags present in MorfFlex might be useful - the given tag would probably strongly predict the presence of corresponding endings (as opposed to derivational affixes and root morphs). These could then be used either for designing specific tagging methods or for enlarging the training data for machine-learning-based taggers.

Secondly, we would like to extend our approaches to a multi-lingual setting. We would especially like to use Universal features included in the Universal Dependencies [18]; these could be also used for more fine-grained morphological analysis in the future. However, there are also many derivational [17] multilingual resources that could be used for the classification of morphs in a similar way to DeriNet.

**Acknowledgements.** This work has been using data, tools and services provided by the LINDAT/CLARIAH-CZ Research Infrastructure (<https://lindat.cz>), supported by the Ministry of Education, Youth and Sports of the Czech Republic (Project No. LM2018101).

We are grateful to Magda Ševčíková and three anonymous reviewers for their helpful comments and suggestions. We thank Šárka Dohnalová for annotating a data sample with classification of morphs.

## References

1. Batsuren, K., et al.: The SIGMORPHON 2022 shared task on morpheme segmentation. In: Proceedings of the 19th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology, pp. 103–116. Association for Computational Linguistics, Seattle, Washington (2022). <https://doi.org/10.18653/v1/2022.sigmorphon-1.11>, <https://aclanthology.org/2022.sigmorphon-1.11>
2. Batsuren, K., Bella, G., Giunchiglia, F.: MorphyNet: a large multilingual database of derivational and inflectional morphology. In: Proceedings of the 18th SIGMORPHON Workshop on Computational Research in Phonetics, Phonology, and Morphology, pp. 39–48. Association for Computational Linguistics (2021). <https://doi.org/10.18653/v1/2021.sigmorphon-1.5>, <https://aclanthology.org/2021.sigmorphon-1.5>
3. Batsuren, K., et al.: UniMorph 4.0: Universal Morphology (2022)
4. Bird, S.: NLTK: the natural language toolkit. In: Proceedings of the COLING/ACL 2006 Interactive Presentation Sessions, pp. 69–72. Association for Computational Linguistics, Sydney, Australia (2006). <https://doi.org/10.3115/1225403.1225421>, <https://aclanthology.org/P06-4018>
5. Bodnár, J., Žabokrtský, Z., Ševčíková, M.: Semi-supervised induction of morpheme boundaries in Czech using a word-formation network. In: Sojka, P., Kopeček, I., Pala, K., Horák, A. (eds.) TSD 2020. LNCS (LNAI), vol. 12284, pp. 189–196. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-58323-1\\_20](https://doi.org/10.1007/978-3-030-58323-1_20)
6. Bolshakova, E.I., Sapin, A.S.: Building a combined morphological model for Russian word forms. In: Burnaev, E., et al. (eds.) AIST 2021. Communications in Computer and Information Science, vol. 1086, pp. 45–55. Springer-Verlag, Berlin (2021). [https://doi.org/10.1007/978-3-030-39575-9\\_16](https://doi.org/10.1007/978-3-030-39575-9_16)
7. Comrie, B., Haspelmath, M., Bickel, B.: The Leipzig glossing rules: conventions for interlinear morpheme-by-morpheme glosses. Max Planck Institute for Evolutionary Anthropology (2008). <https://books.google.cz/books?id=e8B7AQAACAAJ>
8. Donaj, G., Sepesy Maučec, M.: On the use of morpho-syntactic description tags in neural machine translation with small and large training corpora. *Math.* **10**(9) (2022). <https://doi.org/10.3390/math10091608>, <https://www.mdpi.com/2227-7390/10/9/1608>
9. Goldman, O., et al.: The MRL 2022 shared task on multilingual clause-level morphology. In: Proceedings of the The 2nd Workshop on Multi-lingual Representation Learning (MRL), pp. 134–146. Association for Computational Linguistics, Abu Dhabi, United Arab Emirates (Hybrid) (2022). <https://aclanthology.org/2022.mrl-1.14>
10. Goldsmith, J.: Unsupervised learning of the morphology of a natural language. *Comput. Linguist.* **27**(2), 153–198 (2001). <https://doi.org/10.1162/089120101750300490>, <https://aclanthology.org/J01-2001>
11. Hajič, J., et al.: Prague dependency treebank - consolidated 1.0. In: Proceedings of the Twelfth Language Resources and Evaluation Conference, pp. 5208–5218. European Language Resources Association, Marseille, France (2020). <https://aclanthology.org/2020.lrec-1.641>
12. Hajič, J., Hlaváčová, J., Mikulová, M., Straka, M., Štěpánková, B.: MorfFlex CZ 2.0. LINDAT/CLARIN digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University (2020). <http://hdl.handle.net/11234/1-3186>

13. Haspelmath, M.: The morph as a minimal linguistic form. *Morphology* **30**(2), 117–134 (2020). <https://doi.org/10.1007/s11525-020-09355-5>
14. Huang, Z., Xu, W., Yu, K.: Bidirectional LSTM-CRF Models for Sequence Tagging (2015)
15. Ji, D.: Bi-LSTM-CRF (2019). <https://github.com/jidasheng/bi-lstm-crf>
16. Kuznetsova, A., Efremova, T.: Dictionary of Morphemes of the Russian Language. Firebird Publications, Incorporated (1986). <https://books.google.cz/books?id=RPw7AAAACAAJ>
17. Kyjánek, L., Žabokrtský, Z., Ševčíková, M., Vidra, J.: Universal derivations 1.0, a growing collection of harmonised word-formation resources. *Prague Bull. Math. Linguist.* **115**, 5–30 (2020). <https://doi.org/10.14712/00326585.003>
18. Nivre, J., et al.: Universal dependencies v2: an evergrowing multilingual treebank collection. In: Proceedings of the Twelfth Language Resources and Evaluation Conference, pp. 4034–4043. European Language Resources Association, Marseille, France (2020). <https://aclanthology.org/2020.lrec-1.497>
19. Schone, P., Jurafsky, D.: Knowledge-free induction of inflectional morphologies. In: Second Meeting of the North American Chapter of the Association for Computational Linguistics (2001). <https://aclanthology.org/N01-1024>
20. Šiška, Z.: Bázový morfematický slovník češtiny. Univerzita Palackého, Pedagogická fakulta (1998). <https://books.google.cz/books?id=MZMWAAAACAAJ>
21. Slavíčková, E.: Retrográdní morfematický slovník češtiny: s připojenými inventárními slovníky českých morfémů kořenových, prefixálních a sufixálních. Academia (1975)
22. Šojat, K., Srebačić, M., Tadić, M., Pavelić, T.: CroDeriV: a new resource for processing Croatian morphology. In: Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 2014), pp. 3366–3370. European Language Resources Association (ELRA), Reykjavik, Iceland (2014). <http://www.lrec-conf.org/proceedings/lrec2014/pdf/1074.Paper.pdf>
23. Soricut, R., Och, F.: Unsupervised morphology induction using word embeddings. In: Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 1627–1637. Association for Computational Linguistics, Denver, Colorado (2015). <https://doi.org/10.3115/v1/N15-1186>, <https://aclanthology.org/N15-1186>
24. Vidra, J., Žabokrtský, Z., Kyjánek, L., Ševčíková, M., Dohnalová, Š.: DeriNet 2.0. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University (2019). <http://hdl.handle.net/11234/1-2995>
25. Vidra, J., Žabokrtský, Z., Ševčíková, M., Kyjánek, L.: DeriNet 2.0: towards an all-in-one word-formation resource. In: Proceedings of the Second International Workshop on Resources and Tools for Derivational Morphology, pp. 81–89. Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics, Prague, Czechia (2019). <https://aclanthology.org/W19-8510>
26. Žabokrtský, Z., et al.: Towards universal segmentations: UniSegments 1.0. In: Proceedings of the Thirteenth Language Resources and Evaluation Conference, pp. 1137–1149. European Language Resources Association, Marseille, France (2022). <https://aclanthology.org/2022.lrec-1.122>
27. Zeller, B., Šnajder, J., Padó, S.: DERivBase: inducing and evaluating a derivational morphology resource for German. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 1201–1211. Association for Computational Linguistics, Sofia, Bulgaria (2013). <https://aclanthology.org/P13-1118>