Understanding Borrowing through Derivational Morphology: A Case Study of Czech Verbs

Abishek Stephen
Charles University,
Faculty of Mathematics and Physics,
stephen@ufal.mff.cuni.cz

Zdeněk Žabokrtský
Charles University,
Faculty of Mathematics and Physics,
zabokrtsky@ufal.mff.cuni.cz

Abstract

The transfer of morphemes across languages in language contact situations may lead to an alteration in the morphology of the recipient language. One of the possible outcomes can be the introduction of newer word forms or a formation of newer morphological variants of the existing word forms in the recipient language. Languages often borrow nominal roots and morphologically derive them into verbs and are thus integrated into respective derivational classes. This corpus-based analysis for Czech tries to show how synchronic derivational resources can be used to probabilistically analyze the effects of borrowing in language evolution by focusing on morphological integration of the borrowed nominal roots in verb formations.

1 Introduction

In Czech, the use of verbs like *studovat* 'to study', *rezervovat* 'to reserve', *fixovat* 'to fix', *blogovat* 'to blog', and so on is very common. On taking a closer look at such verbs we find that the verbal roots are of foreign origin and not native to Czech. Due to the rich derivational morphology, such verbs in Czech with foreign or borrowed roots take a number of prefixes as well. For example, *reagovat* 'to react' with the addition of the prefix *pře-* becomes *přereagovat* 'to overreact'. Most of these verbs appear within the conjugation class *-ovat*.

According to Blaha (2022), the verbal conjugation with the affix *-ovat* is highly productive for the borrowed (nominal) roots, for especially those that denote an action done using an instrument like *scanovat* 'to scan' or an action defined after the concept denoted by the root like in *investovat* 'to invest' and this pattern has seen an increase in productivity in the last decade. In the presence of multiple conjugation classes in Czech, it looks like for the borrowed roots the affix *-ovat* is the most productive.

The process of integration or allocation of a derivational class happens once the foreign linguistic item is borrowed. In this study, we investigate this phenomenon in more detail to identify the underlying processes of such integration strategies that might have enabled the language to evolve special morphological machinery to deal with the incoming foreign material. Based on the integration strategies, we try to show that the existence of such internal mechanisms renders the language more flexible and competent to accept loanwords.

2 Background and Motivation

Based on sociolinguistic accounts, languages borrow primarily because of a need for a new concept or because of socio-attitudinal reasons (Campbell, 2020). For example, if a language does not have the concept 'fax' it will most probably be borrowed from a language where this concept exists and this borrowing process will be influenced by multiple linguistic and extra-linguistic factors. When a word for a particular concept is already a part of the language's lexicon and still borrowing happens it is for prestige among the other social factors involved. For example, the dominance of Norman French led to the borrowings of culinary vocabulary from French to English even though English had words denoting those concepts (Campbell, 2020).

Languages thus borrow because of the social or attitudinal factors and also for grammatical reasons (Haspelmath, 2009). Ottawa-Hull French speakers might borrow from English because of their preference

Proceedings of the 4th Int. Workshop on Resources and Tools for Derivational Morphology (DeriMo 2023), pages 49–59,
Dubrovnik, Croatia, 5-6 October 2023.

for morphologically simple lexical items over more complex ones in French to express the same referent (Poplack, 2018). In these contact-induced changes, we can find the existence of certain asymmetries in the borrowability of linguistic items (Matras and Sakel, 2007) and these asymmetries reveal the properties of the human language faculty in terms of the stability of linguistic subsystems (Seifart, 2019).

Contact-induced borrowing can occur at variable rates during evolution due to bilingualism, the extent of contact between languages, the typological relatedness of languages or a combination of all of these factors (Thomason, 2001; Nelson-Sathi and List, 2011). According to Mufwene (2001) "Linguistic features are passed on primarily horizontally, more or less on the pattern of features of parasites, through speakers' interactions with members of the same communicative network or of the same speech community. The default condition of linguistic transmission is with modification, however slight this may be. Horizontal and polyploidic transmission independent of generations makes it possible for a new feature to spread fairly rapidly". This transfer of linguistic information can be visualized in parallel to gene transfer in molecular biology. The prokaryotic and eukaryotic evolution shows that the processes through which the gene families are created vary considerably based on the way the genetic material is transferred.

According to Hall et al. (2020) "Horizontal gene transfer (HGT) is particularly prevalent in prokaryotes, where it is one of the main mechanisms contributing to genetic variation and thus evolution". If we were to look into how similar language and genome evolution are then the language evolution may resemble prokaryotic evolution (List et al., 2014). The horizontal or lateral gene transfer begins with the transfer of the foreign DNA in the cytoplasm followed by the recombination into the chromosome and integration with the gene regulatory circuits of the host (Skippington and Ragan, 2013).

Speaking of evolutionary changes, language evolution is usually looked upon in terms of family trees but it has been established that the horizontal components through lexical borrowing also contribute in evolution (Nelson-Sathi and List, 2011; List et al., 2014). Lexical borrowing can replace an existing word, introduce a new word that may co-exist with a native word having the same meaning or it can insert a new word referring to a concept that previously didn't exist in the language (Monaghan and Roberts, 2019).

The incoming lexicon as a result of the lateral transfer or borrowing also needs to be adapted and integrated into the recipient language. This is very similar to the integration of the laterally acquired foreign genetic material into a host cell. According to Filipović (1981), the adaptation of loanwords on the morphological level is concerned primarily with the formation of its citation form. And this analysis is made based on the transmorphemization[1]. Other claims suggest that speakers integrate verbs merely as lexical labels while others use them, to various degrees, as predicate-initiating devices (Matras and Adamou, 2020).

In case of the presence of multiple conjugation classes, for example, in Czech the citation forms of the verbs can take the affixes[2] *-ovat*, *-it*, *-at*, *-nout*, and so on . In other languages too, we see similar patterns. There has to be one form that is easily accessible or has a higher combinatory potential and thus will get attached to the incoming foreign root readily. In other words, a language could have possibly evolved or devised a mechanism for handling the morphology of the foreign linguistic materials by spreading the existing morphological processes to the borrowed vocabulary. In Croatian[3], for the loanverbs, the English root gets attached to the infinitive affixes *-irati*, *-avati*, *-ivati*, *-ovati*, *-ati* according to the Croatian morphology and we find verbs like *intervjuirati* 'to interview', *flertovati* 'to flirt', etc. In Poplack (2018) we observe that the English bare infinitive itself serves as the root for conjugation when incorporated into Quebec French. English-origin verbs are assimilated into the *-er* group and conjugated according to French morphology like the verbs *mover* 'to move', *runner* 'to run', *shopper* 'to shop' and *skipper* 'to skip' to name a few.

Such evidences show that a language *somehow* assigns a conjugation class to the verbs formed with borrowed roots (Figure 1). Out of the total collection of roots or the *bag of roots* in the lexicon, verb

[1]According to Filipović (1980), transmorphemization is one of the forms of substitution that comprises all the changes appearing in the adaptation of bound morphemes as they pass from the donor language to the recipient language.

[2]For simplicity we speak about affixes, but in fact the presented strings contain also endings (or also other affixes).

[3]As informed by Matea Filko through personal communication.

formation happens based on some underlying combinatorial mechanism let alone the conditions placed by different components of grammar like phonology, syntax, semantics, etc. With the current study, we empirically explore the possible reasons why in Czech the verbs with a borrowed root almost always fall into the *-ovat* conjugation class. We reason as to why only this particular affix is preferred over the others and how a language decides upon such a selection. We assume that these derivational processes make it possible for a language to accept foreign linguistic units and try to explore the reasons of morphological integration based on corpus analysis.

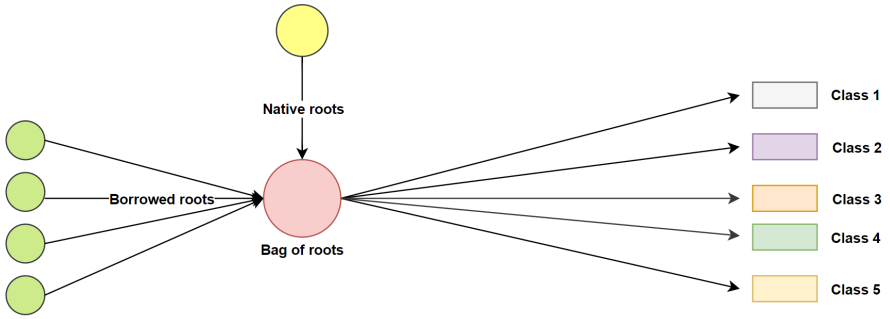

Figure 1: A potential flow of the borrowed and native roots into different conjugation classes

3 Approach

For our analysis we use DeriNet (Vidra et al., 2019). DeriNet[4] is a lexical network that models word-formation relations in the lexicon of Czech. We only take into consideration the loanverbs i.e. verbs with a borrowed stem such as in Figure 2 based on the conjugation classes. There are tags within DeriNet for loanwords which were extracted in a supervised manner from multiple corpora based on language specific rewrite rules. Out of all the loanverbs we only consider those that are not idiosyncratic and are attested in the corpora. We present the unigram frequencies and the relative frequencies of native verbs and loanverbs belonging to different conjugation classes. We also calculate the conditional probabilities and entropies for the distribution of verbs with native and borrowed roots. We additionally calculate Dice coefficient and some other relevant statistic measures. The major topic of investigation is that when the verbs are borrowed into a language, there are certain affixes sensitive to get attached to the incoming foreign root. This when viewed through the lens of loanword integration and adaption seems like a rather probabilistic process than a discrete one. One of the possible reasons could be that the languages, in our case Czech could have possibly evolved special mechanisms to incorporate foreign linguistic material.

Frequency effects in this regard have also gained quite an attention. Pagel et al. (2007) report that the higher-frequency words are more stable and resistant to change or evolution. Such a word form is less likely to be replaced and it also won't admit co-existence with a semantically congruent counterpart but if the word form is represented in a less robust fashion then it is more likely to be replaced or to admit co-existence with a borrowed word form (Monaghan and Roberts, 2019). Hence, we find it worthwhile to analyze frequency effects in this regard i.e. how likely is it that a highly frequent word form would be a borrowed word, and so on. We also compare the derivational rate for the verbs with native and borrowed stems. The frequency of derivational nodes could possibly shed light on the difference in morphological productivity of both the classes of verbs under investigation.

4 Evaluation and Results

For our experiments, we consider only the verbs in DeriNet. We take the native and loanverbs based on corpus attestations (Table 1) i.e. the unigram corpus frequencies of the verbs belonging to both groups must be greater than 1 in the corpora. The absolute frequencies in DeriNet are taken from the Czech National Corpus, SYNv4 (Křen et al., 2016). As we focus on the loanverbs, we analyzed the derivational

[4] https://ufal.mff.cuni.cz/derinet

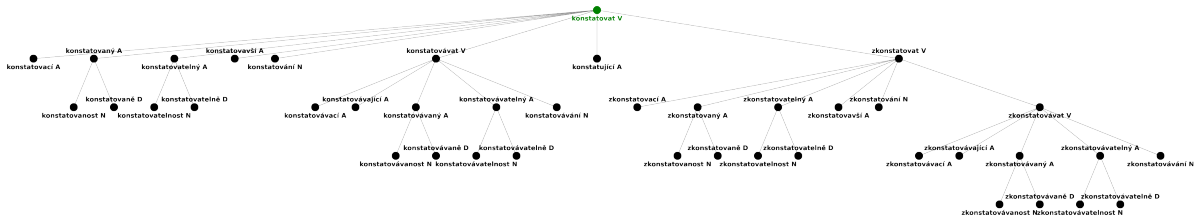

Figure 2: Word-formation relations for the most frequent loanverb *konstatovat* in DeriNet

affixes only for verbs and not for other parts-of-speech. On calculating the unigram frequencies for the

Type	Total verbs	Corpus attested
Native root	42930	19854
Borrowed root	13378	3972

Table 1: Frequencies of verbs in DeriNet

corpus-attested verbs in the DeriNet data we found that most of the verbs with a native root have the derivational affix *-at* followed by the affixes *-it* and *-ovat*. But we also find that a limited set of verbs with the affixes *-ít* and *-et* occur almost as frequently as the verbs with affixes *-at*, *-it* and *-ovat*. For the verbs with a borrowed root, almost all the verbs have the derivational affix *-ovat* followed by *-ovávat* and *-it*. (Table 2). Since our focus is on the loanverbs, we compare the frequencies only with those affixes with the native roots for which there are adequate counterparts with the borrowed roots. The frequencies point out that the conjugation classes of verbs have different *choices* of roots. We also compute the entropies for the distributions of the verbs with native and borrowed roots. Following the standard notion of entropy, we compute entropy *H* of a particular affix *X* as the negative summation of the log of relative frequencies of the affixes *x*, within the group of verbs with borrowed or native roots, *P(x)*.

$$H(X) = -\sum_{x \in X} P(x) log_2 P(x) \tag{1}$$

To check for the strength of attraction between the conjugation affixes and the type of roots Dice coefficient was used as a statistical measure. It is one of the most common association measures used to detect collocations. Dice coefficient outperforms other association measures like mutual information, etc in the task of collocation detection (Kolesnikova, 2016). But for our analysis, we assume the combination of the root and affix is equivalent to a collocation.

Affixes	Native root	Tokens	Borrowed root	Tokens
-at	6481	19225633	62	10724
-it	4492	25400609	141	130751
-ovat	4132	10791009	3377	3749598
-nout	1780	6024365	41	22712
-ovávat	413	128614	307	3307
-et	778	12355416	3	347
-ět	519	6481870	28	860
-át	67	5676169	0	0
-ít	195	11439639	0	0
-ýt	28	228658	0	0

Table 2: Frequencies of derivational affixes

Origin of root	Affix	Lexicon frequencies		Corpus frequencies	
		P(Suffix|Origin)	Entropy	P(Suffix|Origin)	Entropy
Native	*-at*	0.326	2.330	0.200	2.803
	-it	0.230		0.254	
	-ovat	0.208		0.108	
	-nout	0.090		0.060	
	-ovávat	0.020		0.001	
	-et	0.040		0.123	
	-ět	0.026		0.064	
	-át	0.003		0.057	
	-ít	0.009		0.114	
	-ýt	0.001		0.002	
Borrowed	*-at*	0.015	0.873	0.023	0.401
	-it	0.035		0.033	
	-ovat	0.850		0.958	
	-nout	0.010		0.006	
	-ovávat	0.078		0.001	
	-et	0.001		0.000	
	-ět	0.007		0.000	
	-át	0.000		0.000	
	-ít	0.000		0.000	
	-ýt	0.000		0.000	

Table 3: Probabilities and entropies of derivational affixes

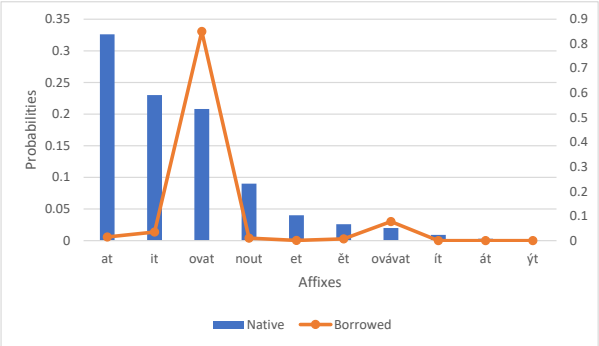

Figure 3: Probabilities based on lexicon frequencies

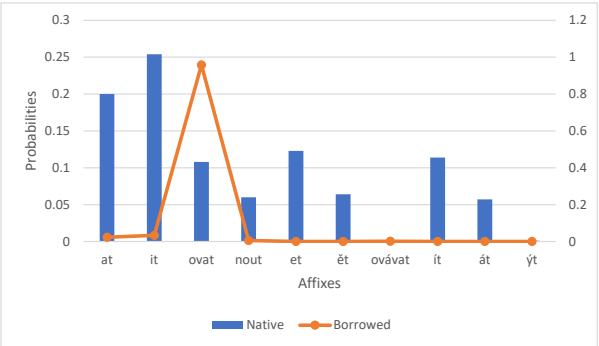

Figure 4: Probabilities based on corpus frequencies

	-at	*-it*	*-ovat*	*-nout*	*-ovávat*	*-et*	*-ět*	*-át*	*-ít*	*-ýt*
Native root	0.322	0.404	0.188	0.114	0.002	0.220	0.122	0.107	0.205	0.005
Borrowed root	0.001	0.009	0.406	0.005	0.001	0	0	0	0	0

Table 4: Dice coefficient based on corpus frequencies

In Table 3, we observe that the entropy for the conjugation classes of verbs with native roots is much higher than the distribution of verbs with borrowed roots. The trend that the entropy of affixes is lower with borrowed roots is even stronger in a corpus with running text frequencies, as opposed to lexicon frequencies. One of the reasons is that the borrowed root almost always occurs with the affix *-ovat* (Figure 3 and Figure 4). This can be viewed as an analogy of the form a : b = c : x. In historical

	do-	*roz-*	*o-*	*po-*	*pod-*	*od-*	*u-*	*v-*	*vy-*	*z-*	*za-*	*pře-*	*před-*
Native root	616	649	568	661	134	506	652	623	1228	1263	1226	454	117
Borrowed root	108	38	70	44	7	104	23	47	187	338	214	140	14

Table 5: Lexicon frequencies of the prefixed verbs with native and borrowed roots

linguistics, an analogical change can be defined as a process whereby one form of language becomes more like another with which it is somehow associated (Arlotto, 1972). The analogy is also referred to as internal borrowing where a language *borrows* some of its own patterns to change other patterns (Campbell, 2020). The conditional probabilities do indicate that one of the patterns of the derivation of verbs with a native root has been applied to derive verbs with a borrowed root i.e. the conjugation class *-ovat*.

Furthermore, the Dice coefficient scores in Table 4 also support that the affix most *sensitive* to borrowed roots is *-ovat*. This behaviour could be attributed to a quicker processing of verbs with this affix. Assuming the lexical units that have a higher information load are more costly to process, the lexical processing cost becomes directly proportional to the amount of information. The conditional probabilities in Table 3 indicate that the verbs with the affix *-ovat* carry the least amount of information[5] and hence they are easier to process as compared to the other verbs with different affixes.

The argument around lexical processing itself requires its own space of discussion which is beyond the scope of this paper. But we would like to examine if the length of the affix plays any role behind the specific selection of *-ovat* for the borrowed roots. Most the affixes are of length 2 i.e. *-at*, *-it*, *-et*, and so on followed by the affixes *-ovat* and *-ovávat* with lengths 4 and 6 respectively. In word recognition and recall tasks, immediate memory span is better with short than with long words (Baddeley et al., 1975). The weighted average of the length of the affix and the conditional probabilities based on corpus frequencies were calculated (see Table 3) and it was found that the average length of the conjugation class affix is 2.3 for the verbs with the native roots and for the verbs with the borrowed roots it is 3.9. This again falls in accordance with the most preferred affixes by the both the type of roots. It is difficult to say if the borrowed roots fall into the *-ovat* class and hence a longer affix is preferred or it is the other way round. In Croatian, it can be speculated that the borrowed stems take the conjugation class with a longer affix like in *intervjuirati* 'to interview'. In Slovak, we find examples like *fotografovať* 'to take pictures' and also Polish *komentować* 'to comment'. Based on these examples, we might reach a probable conclusion that the verbs derived using a borrowed stem is *marked* with a longer suffix in the presence of multiple conjugation classes where the affix lengths vary. It might also indicates that the speakers of these languages label the loanverbs with a longer affix almost always but since we only deal with Czech primarily in this study, we do not make any concrete claims about other languages.

We also investigate if the prefixes play any role in the integration strategies of the borrowed roots. In Table 5, we present the lexicon frequencies of the prefixed verbs with native and borrowed roots. The difference in both the classes do not present any striking contrast. Moreover, based on the derivational trees (see Appendix) we can infer that the formation of verbs begins with the combination or selection of a root and a conjugation affix which is then followed by derivations by the addition of the prefixes and the roots do not seem to play any significant role in the selection of the prefixes. But in any case, we did calculate some frequency measures (Table 6) for the verbs with the conjugation affix *-ovat* and found that nearly 70% of the verbs with a native root are prefixed and only 45% of the verbs with a borrowed root are prefixed. To analyze the morphological productivity[6], we calculated the average number of derivational nodes in the derivational tree for verbs with native and borrowed roots present in DeriNet (Figure 5). The results show that on an average a verb with a native root has 34.7 derived word forms whereas a verb with a borrowed root has 35.1 derived word forms. This indicates that most of the derivational processes are similar for the native and borrowed words.

[5]The amount of information carried is the negative logarithm of the probability.
[6]In a narrow sense the conditional probabilities in Table 3 can also serve as an indicator of morphological productivity.

	do-	*roz-*	*o-*	*po-*	*pod-*	*od-*	*u-*	*v-*	*vy-*	*z-*	*za-*	*pře-*	*před-*
Native root	112	123	141	129	30	126	129	106	306	270	293	119	26
Borrowed root	69	23	51	37	4	66	25	23	142	291	160	91	11

Table 6: Lexicon frequencies of the prefixed verbs with native and borrowed roots with affix *-ovat*

```
if word.is_loanword != "None" and word.absolute_count > 1:
   if word.is_loanword == "True":
           count_loan++
           count_children_loan += word.get_all_children()
           if word.pos == "VERB":
               count_verb_loan++
               count_children_verb_loan += word.get_all_children()
   else:
           count_native++
           count_children_native += word.get_all_children()
           if word.pos == "VERB":
               count_verb_native++
               count_children_verb_native += word.get_all_children()

avg_children_verb_loan: float = count_children_verb_loan / count_verb_loan
avg_children_verb_native: float = count_children_verb_native / count_verb_native
```

Figure 5: Pseudocode for extracting the number of derivations per word from DeriNet

For English as a donor language, Monaghan and Roberts (2019) report that for the mid- to high-frequency words in English the likelihood of borrowing drops but for mid- to low-frequency words (with frequencies less than one per ten thousand) the relationship is positive and monotonic i.e. the likelihood of borrowing increases. For analyzing the distribution of the verbs with borrowed and native roots, considering Czech as a recipient language, we calculated the probability of finding a loanword and compared it with the probability of finding a verb with a borrowed stem (Figure 6 and Figure 7). We observe that locating a loanword or a loanverb increases with an increase in the corpus frequencies i.e. a loanword can be as highly frequent as a native word.

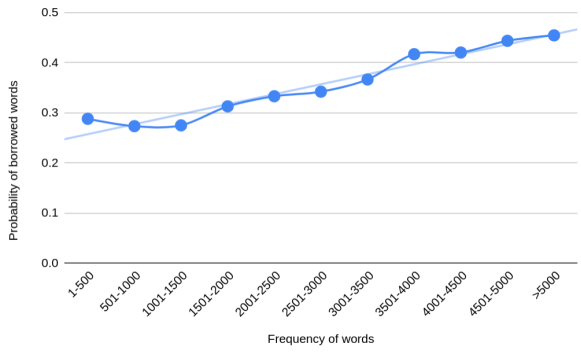

Figure 6: Probability of locating a loanword

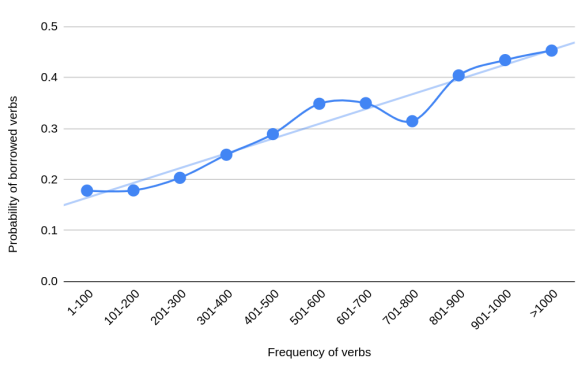

Figure 7: Probability of locating a verb with borrowed root

Based on all of the above statistical measures, we can conclude that the integration strategies in Czech treat the borrowed linguistic material in a very similar manner like the native vocabulary. The

derivations proceed in a very similar direction as the average number of derived word-forms indicates. It seems like due to the presence of multiple conjugation classes, loanverbs are preferably conjugated using the class with the longer affix. The observations based on the corpus frequencies indicate the *-ovat* conjugation class is neither the most frequent nor the least frequent choice of the native verbs. The aim of a morphological system is not to increase chaos and thus it identifies and expands the recurrent patterns to the borrowed words. This process also characterizes the cognitive capacity in a narrow sense.

For handling the integration, we assume that Czech chooses the pattern that has a central tendency given that a corpus is statistically dispersed. The measures of central tendency can be used to summarize the profile of verbs with either type of roots. We already know the probabilities and the corpus frequencies show that the native verbs falling within the *-ovat* conjugation class are neither the most frequent nor the least frequent. They lie somewhere in the middle of the distribution. The median of the corpus frequencies of verbs with a native root in Table 2 happens to be 8.6 million which is close to the corpus frequency of the conjugation class *-ovat*. There seems to a higher probability that the choice of the conjugation class for loanverbs should fit around the median so as to keep the morphological system out of chaos. It is difficult to generalize this behaviour due to the lack of comparative corpus analysis across a good number of languages but it does seem to be prospective. The findings are purely empirical. There is a possibility that some extra-linguistic factor initiated the assimilation of loanverbs into the *-ovat* class. Language contact situations are complex and hence we cannot rule out the possibility that the integration strategies can be influenced by other factors as well.

5 Conclusion

This study analyzes the loanverbs in Czech based on DeriNet. The corpus analysis showed that the loanverbs almost always fall into the conjugation class *-ovat*. This can be seen as a strategy to mark the loanverbs with a longer suffix to indicate that the root is borrowed. Other statistical measures indicate that to keep the morphological system out of chaos that can be caused due to the incoming borrowed words, the central derivational process is extended towards handling the morphology of loanwords in the presence of multiple verb conjugation classes. These underlying mechanisms act as a positive pressure for accepting borrowings and thus contribute to the evolution of language in terms of its vocabulary range and morphological specializations to name a few among the various other modifications. Thus, the verb integration strategies or in this study the derivational processes led by the conjugation classes play a vital role in language change and evolution over time.

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 thank the anonymous reviewers for their valuable comments and Michal Olbrich for his technical support with DeriNet.

References

Anthony T. Arlotto. 1972. Introduction to Historical Linguistics. Houghton Mifflin, Boston.

Alan D. Baddeley, Neil Thomson, and Mary Buchanan. 1975. Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior* 14(6):575–589.

Ondrej Blaha. 2022. Dynamics of Conjugation Pattern "Kupuje" in Contemporary Czech (On Material of Journalistic Texts, 1990-2019). *Bohemica Olomucensia* 14(2):40–54.

Lyle Campbell. 2020. *Historical Linguistics: An Introduction*. Edinburgh University Press.

Rudolf Filipović. 1980. *Transmorphemization: Substitution on the Morphological Level Reinterpreted*, Revue publiée par les Sections romane, italienne et anglaise de la Faculté des Lettres de l'Université de Zagreb, volume 25 (-), pages 1–8.

Rudolf Filipović. 1981. *Morphological Categories in Linguistic Borrowing*, Revue publiée par les Sections romane, italienne et anglaise de la Faculté des Lettres de l'Université de Zagreb, volume 26 (-), pages 197–207.

Rebecca J. Hall, Fiona J. Whelan, James O. McInerney, Yaqing Ou, and Maria Rosa Domingo-Sananes. 2020. Horizontal Gene Transfer as a Source of Conflict and Cooperation in Prokaryotes. *Frontiers in Microbiology* 11.

Martin Haspelmath. 2009. *II. Lexical borrowing: Concepts and issues*, De Gruyter Mouton, Berlin, New York, pages 35–54.

Olga Kolesnikova. 2016. Survey of Word Co-occurrence Measures for Collocation Detection. *Computacion y Sistemas* 20:327–344.

Michal Křen, Václav Cvrček, Tomáš Čapka, Anna Čermáková, Milena Hnátková, Lucie Chlumská, Tomáš Jelínek, Dominika Kováříková, Vladimír Petkevič, Pavel Procházka, Hana Skoumalová, Michal Škrabal, Petr Truneček, Pavel Vondřička, and Adrian Zasina. 2016. SYN v4: large corpus of written czech. LINDAT/CLARIAH-CZ digital library at the Institute of Formal and Applied Linguistics (ÚFAL), Faculty of Mathematics and Physics, Charles University.

Johann-Mattis List, Shijulal Nelson-Sathi, Hans Geisler, and William Martin. 2014. Networks of lexical borrowing and lateral gene transfer in language and genome evolution. *BioEssays* 36(2):141–150.

Yaron Matras and Evangelia Adamou. 2020. *Borrowing*, Routledge, pages 237–251.

Yaron Matras and Janette Sakel. 2007. Investigating the mechanisms of pattern replication in language convergence. *Studies in Language* 31(4):829–865.

Padraic Monaghan and Seán G. Roberts. 2019. Cognitive influences in language evolution: Psycholinguistic predictors of loan word borrowing. *Cognition* 186:147–158.

Salikoko S. Mufwene. 2001. *Language contact, evolution, and death: how ecology rolls the dice*, Cambridge University Press, page 145–166. Cambridge Approaches to Language Contact.

Shijulal Nelson-Sathi and Johann-Mattis List. 2011. Networks uncover hidden lexical borrowing in Indo-European language evolution. *Proceedings of the Royal Society* pages 1794–1803.

Mark Pagel, Quentin Atkinson, and Andrew Meade. 2007. Frequency of Word-Use Predicts Rates of Lexical Evolution throughout Indo-European History. *Nature* 449:717–20.

Shana Poplack. 2018. Borrowing in the speech community. In *Borrowing: Loanwords in the Speech Community and in the Grammar*, Oxford University Press.

Frank Seifart. 2019. *2. Contact-induced change*, De Gruyter Mouton, Berlin, Boston, pages 13–23.

Elizabeth Skippington and Mark A. Ragan. 2013. *Lateral Genetic Transfer and Cellular Networks*, Springer New York, New York, NY, pages 123–135.

S.G. Thomason. 2001. *Language Contact*. Edinburgh University Press.

Jonáš Vidra, Zdeněk Žabokrtský, Magda Ševčíková, and Lukáš Kyjánek. 2019. 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*. Charles University, Faculty of Mathematics and Physics, Institute of Formal and Applied Linguistics, Prague, Czechia, pages 81–89.

6 Appendix

Figure 8 and Figure 9 show the word-formation relations for the loanverbs *parkovat* 'to park' and *komentovat* 'to comment'. The derivational trees presented here only contain few sub-branches and nodes that indicate the prefixation of the verbs. The complete visualization of the trees can be viewed using the DeriNet online viewer available at: https://ufal.mff.cuni.cz/derinet/derinet-viewer. Table 7 contains the corpus frequencies of the top 50 most frequent native verbs and loanverbs.

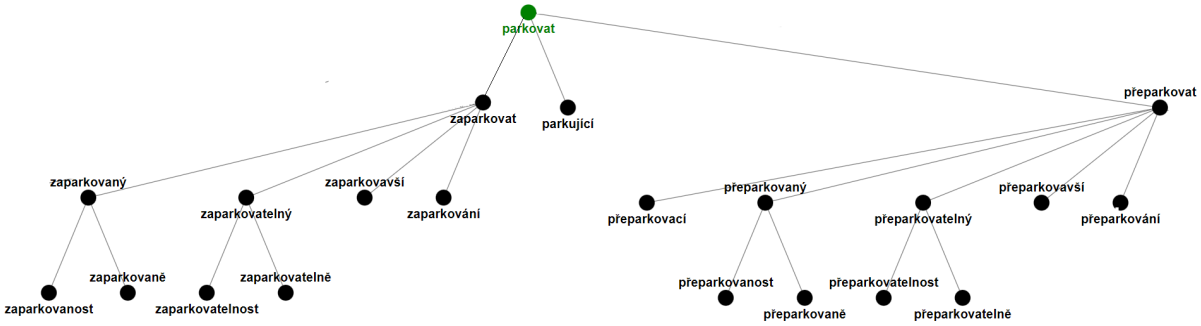

Figure 8: Word-formation relations for the loanverb *parkovat* in Derinet

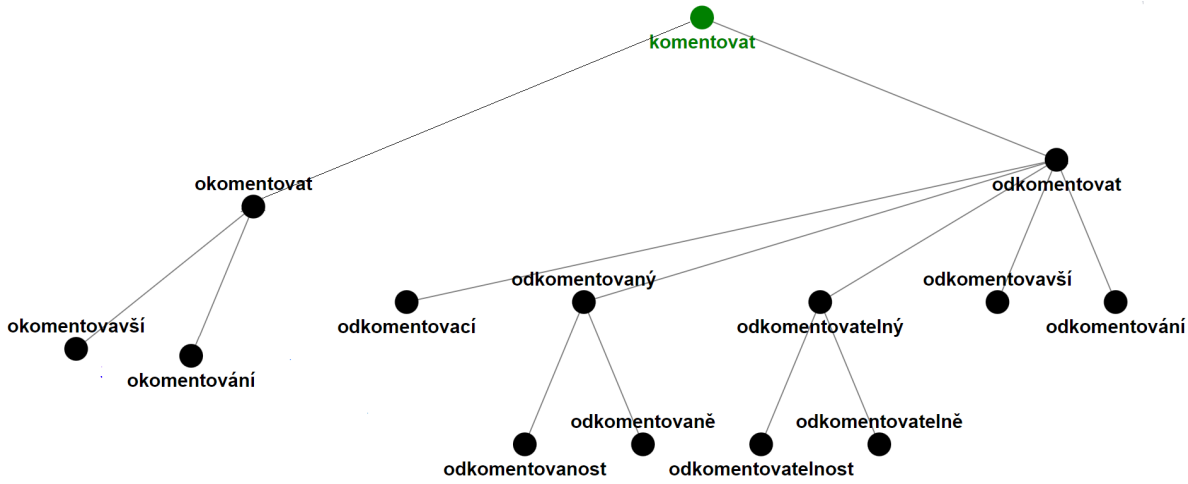

Figure 9: Word-formation relations for the loanverb *komentovat* in Derinet

Table 7: Corpus frequencies of 50 most frequent verbs with native and borrowed roots in DeriNet

Native	Corpus frequencies	Borrowed	Corpus frequencies
muset	6758539	konstatovat	452229
chtít	6478359	investovat	272892
hrát	3334886	komentovat	248282
vědět	2973377	fandit	86867
potvrdit	996252	kontaktovat	83964
mluvit	924879	nominovat	78456
dávat	883624	instalovat	77583
jezdit	882479	rezignovat	72944
využít	875927	charakterizovat	72458
hledat	842844	kombinovat	67995
věnovat	828345	argumentovat	67867
řešit	825438	zareagovat	66842
přijet	798967	testovat	66225
vybrat	726762	zaregistrovat	65575
dosáhnout	724422	angažovat	59529
umět	674523	produkovat	58693
držet	664526	definovat	58582
nabídnout	663885	konkurovat	56473
psát	654944	akceptovat	53563
zajistit	653733	aplikovat	49925
představovat	648567	rekonstruovat	48446
připravovat	645968	parkovat	47738
koupit	628786	blokovat	46552
prohrát	592736	identifikovat	42783
odejít	578272	kopírovat	39466
bývat	577669	režírovat	38935
pořádat	553959	stabilizovat	38297
podívat	529262	nastudovat	37494
ztratit	525653	sportovat	37454
zúčastnit	522742	zrekonstruovat	36972
projít	522254	fotografovat	35896
odehrát	498576	eliminovat	35856
pohybovat	488446	iniciovat	34688
oznámit	488398	evakuovat	33585
zajímat	479852	avizovat	29564
uvidět	477494	formulovat	27853
vystoupit	469273	deklarovat	27839
upozornit	468482	kompenzovat	27287
bránit	466535	stagnovat	26643
sedět	452492	zkolabovat	25469
navštívit	448659	emigrovat	24536
vycházet	436446	interpretovat	23824
odpovídat	435243	finišovat	22958
popsat	435226	natankovat	9835
sejít	434936	pogratulovat	9742
připomínat	434872	proinvestovat	9726
připomenout	433386	marodit	9488
plánovat	433345	konkretizovat	9436
vzpomínat	428383	restaurovat	9355
určit	427396	zkorigovat	8988