
Boris Kern

CONSIDERING WORD FORMATION IN COMPILING DICTIONARIES

Abstract In this article, we aim to illustrate, using the example of Slovenian, which aspects need to be considered when compiling a new dictionary. In compiling dictionaries, considering word formation is crucial in several phases: a) selecting criteria for including words in the dictionary, b) uniformity in dictionary definitions of headwords that have some word-formation relation, and c) consistency in qualifying certain senses as stylistically marked. Regarding the criteria for including words in the dictionary, this article pays special attention to statistical analysis of the correlation between multistage derivatives and the frequency of individual words in word-formation chains, in corpus sources, and in the case at hand the Gigafida 2.0 and metaFida corpora. It is assumed that the frequency of a word decreases as its formation stage becomes higher. This article presents dictionary solutions as indicated by dictionary concepts for the monolingual dictionary of standard Slovenian.

Keywords word formation; multistage word formation; Slovenian lexicography

1. Introduction

Like other Slavic languages, Slovenian lexemes are characterized by an extremely rich morphemic structure, which is a result of multistage word formation. For example, in the first stage, the adjective *mlad* ‘young’ yields the noun *mladost* ‘youth’, which in turn yields the adjective *mladosten* ‘youthful’ in the second stage, which in turn yields the noun *mladostnik* ‘adolescent’ in the third stage, which in turn yields the possessive adjective *mladostnikov* ‘adolescent’s’ in the fourth stage (Skarżyński, 1999; 2003; Olejniczak, 2003; Kern, 2010; 2017; Vidovič Muha, 2018).¹ This linguistic fact influences certain aspects of compiling dictionaries for Slavic languages at several stages:² a) selecting criteria for including words in the dictionary, b) uniformity in dictionary definitions of headwords that have some word-formation relation, and c) consistency in qualifying certain senses as stylistically marked (Skarżyński, 1999; 2003; Olejniczak, 2003; Burkacka, 2012; Kern, 2010; 2017). Regarding the criteria for including words in the dictionary, it is necessary to define strategies for cases

¹ From 2021 to 2024, the project Formant Combinatorics in Slovenian is being carried out in collaboration with ZRC SAZU, the Jožef Stefan Institute, and the University of Maribor. The objective of the project is to explore the combinatorics of word-formation formants, which will make it possible to present the characteristics of word formation and semantic extension mechanisms of Slovenian on contemporary language material. The language technology objective of the project is a pioneering creation of the first training set and the first language technology application allowing automatic morpheme segmentation of Slovenian words (see <https://isjfr.zrc-sazu.si/en/programi-in-projekti/formant-combinatorics-in-slovenian>).

² It is worth noting that from the sixteenth to eighteenth century—the period when efforts to compile monolingual dictionaries emerged—the “word-family” arrangements of lexemes in dictionaries was a serious competitor to the alphabetical arrangement of lexemes (August, 1993, pp. 188–189).

in which so-called word-formation gaps or hypothetical derivations³ occur in the word-formation chains. A particular element of a chain may not be present in the language. For example, in the Slovenian word-formation chain *okus* ‘taste’ → *okusiti* ‘to taste’ *pf.* → *okuševati* ‘to taste’ *impf.* → *okuševalec* ‘taster’, the verb form *okuševati* occurs only once in the Slovenian reference corpus Gigafida 2.0. For this reason, it is not included in dictionaries of Slovenian. Hypothetical derivatives can also be observed at the end of word-formation chains, where, according to previous studies, diminutives, possessive adjectives, adjectives ending in *-ni* or *-ski/-ški*, nouns ending in *-ost*, and verbal nouns (e.g., *zaspanost* ‘sleepiness’) occur most frequently in Slovenian. Regarding qualifying the senses as stylistically marked, it should be added that it is of course neither necessary nor usual for all words in a word-formation chain to have the same connotative meaning. However, it is useful in lexicographic work to review all multistage derivatives, both in terms of qualification and dictionary definitions.

In dictionary definitions, it is sensible to use unified classificatory (*genus proximum*) and distinctive semantic components (*differentia specifica*) within word-formation chains where possible (e.g., **vegán** kdor ne je hrane živalskega izvora in zavrača vsakršno izkoriščanje živali → **vegánka** ženska, ki ne je hrane živalskega izvora in zavrača vsakršno izkoriščanje živali → **vegánstvo** prehranjevanje s hrano, ki ni živalskega izvora, in zavračanje vsakršnega izkoriščanja živali; **vegan** someone who does not consume animal-based food and rejects any form of animal exploitation → **female vegan** a woman who does not consume animal-based food and rejects any form of animal exploitation → **veganism**: eating food that is not of animal origin and rejecting any form of animal exploitation’).⁴ It is also important to maintain a unified semantic structure, in terms of the relationship between the primary meaning and subordinate meaning. An example of inconsistency can be seen with the verbs *videti* ‘to see,’ *slišati* ‘to hear,’ and *poslušati* ‘to listen’ in the Dictionary of the Slovene Standard Language, 2nd edition (2024): in the second meaning of *videti* ‘to see’ and *slišati* ‘to hear,’ both express ‘becoming aware through sight or hearing,’ whereas the verb *poslušati* ‘to listen’ has the same semantic transition into the field of mental perception expressed in a subordinate meaning (‘to perceive and understand the content of something by listening’). These three examples are systematically similar, so the semantic differentiation should be unified, especially in cases involving derived terms.

A methodology that seems to be particularly useful for analyzing such material, in all three phases listed above, is an analysis based on multistage word formation. This methodology extends the binary relation between the motivating and motivated word to larger groups of words, which include both direct derivatives (e.g., the noun *mladost* ‘youth’ with regard to the adjective *mlad* ‘young’) and indirect derivatives (e.g., the possessive adjective *mladostnikov* ‘adolescent’s’ with regard to the noun *mladost* ‘youth’).⁵

³ The hypothetical derivative is a derivative whose introduction within word-formation chain is necessary especially in order to explain a higher-stage derivative that is attested in use.

⁴ Examples are from the eSSKJ: Dictionary of the Slovene Standard Language (2016–).

⁵ In this context, the studies by Čibej (2021) and Dobrovoljc et al. (2017) are also relevant for Slovenian.

2. Word-Formation Aspects in Dictionary Concepts for Slovenian

In Slovenian linguistics, two dictionary concepts were published approximately simultaneously: the first was a proposal for a modern Slovenian dictionary (Krek et al., 2013; hereinafter: Concept I), followed by a concept for a new explanatory dictionary of standard Slovenian by lexicographers at the Lexicological Section of the ZRC SAZU Fran Ramovš Institute of the Slovenian Language (Gliha Komac et al., 2015; hereinafter: Concept II).

2.1 Concept I: A Modern Slovenian Dictionary

Concept I undoubtedly presents a well-thought-out approach to the dictionary in a manner that follows modern lexicographic trends. However, the concept pays relatively little attention to the presentation of word-formation data. It mentions “word formation–related forms,” noting that it is necessary to consider how to address the problem of word formation–related forms in relation to polysemy (e.g., in the case of diminutives; Krek et al., 2013, p. 73), or the problem of asymmetry between formation–indicated diminutiveness and semantic non–diminutiveness, such as *jopica* ‘small sweatshirt’ versus *marelica* ‘apricot’ (and not ‘small umbrella’; Krek et al., 2013, pp. 74–75). The concept also states that the dictionary lists words as independent headwords when they feature competing word–formation suffixes that have had different normative values in previous linguistic resources (*-lec* vs. *-vec*; *zajedavec* vs. *zajedalec* ‘parasite’). The principle accepted is that, if both forms exceed the frequency threshold for inclusion in the lexical database, the words are treated as independent headwords “since it is possible to anticipate both semantic differentiation and variation in textual context and usage, which is usually conditioned by the frequency of each form.” The concept also mentions that the lexical database does not treat prefixes such as *evro-*, *pro-*, *anti-*, *e-*, and so on as independent headwords,⁶ but either as standalone headwords (e.g., *evro* ‘euro’) or as whole words to which such prefixes are attached (e.g., *biomasa* ‘biomass’; Krek et al., 2013, p. 75). The dictionary was not realized over the following years; however, due to the concept’s innovativeness, it undoubtedly deserves to be included in this research.

2.2 Concept II: A New Explanatory Dictionary

On the other hand Concept II, which was approved and adopted in 2015 by the Research Council of the ZRC SAZU Fran Ramovš Institute of the Slovenian Language, the ZRC SAZU Research Council, the ZRC SAZU Class of Philological and Literary Sciences, and the SAZU Presidency (eSSKJ 2017, p. 1), is a basis for the third edition of eSSKJ *Slovar slovenskega knjižnega jezika* (Standard Slovenian Dictionary, hereinafter: eSSKJ),⁷ which has been available as a growing dictionary⁸ on the lexicographic portal Fran since 2016. It is a new informative–normative explanatory dictionary of standard

⁶ This was the practice in some other Slovenian dictionaries: *Slovar slovenskega knjižnega jezika* (Standard Slovenian Dictionary, 1970–1991, 2014) and *Slovenski pravopis* (Slovenian Normative Dictionary, 2001).

⁷ The first edition was published between 1970 and 1991, and the second in 2014.

⁸ New dictionary entries are added each year; the exception was 2020.

Slovenian, with dictionary entries created entirely anew based on the analysis of contemporary language material. eSSKJ is the central lexicographic project of the ZRC SAZU Fran Ramovš Institute of the Slovenian Language and is planned to be a dictionary with approximately 100,000 dictionary entries, introduced at the macrostructural level by single-word headwords (Gliha Komac et al., 2015, p. 3). So far, 3,432 entries have been published in the dictionary.

The eSSKJ dictionary includes a word-formation tab, which, according to Concept II, is shown only for noun entries. In this tab, possessive adjectives, diminutives, feminatives, and masculatives are listed with hyperlinks if they meet frequency criteria and thus appear as independent entries.⁹ The terms feminative and masculative are used in a broad sense, encompassing examples of parallel derivatives (e.g., *igralec* ‘actor’ vs. *igralka* ‘actress’),¹⁰ but the terms themselves are not listed in the tab; instead, they are replaced by the labels *ženska oseba* ‘female’ or *moška oseba* ‘male’ before the related entry (see Figure 1). Feminatives are listed for paired masculine nouns on Fran, but not vice versa.

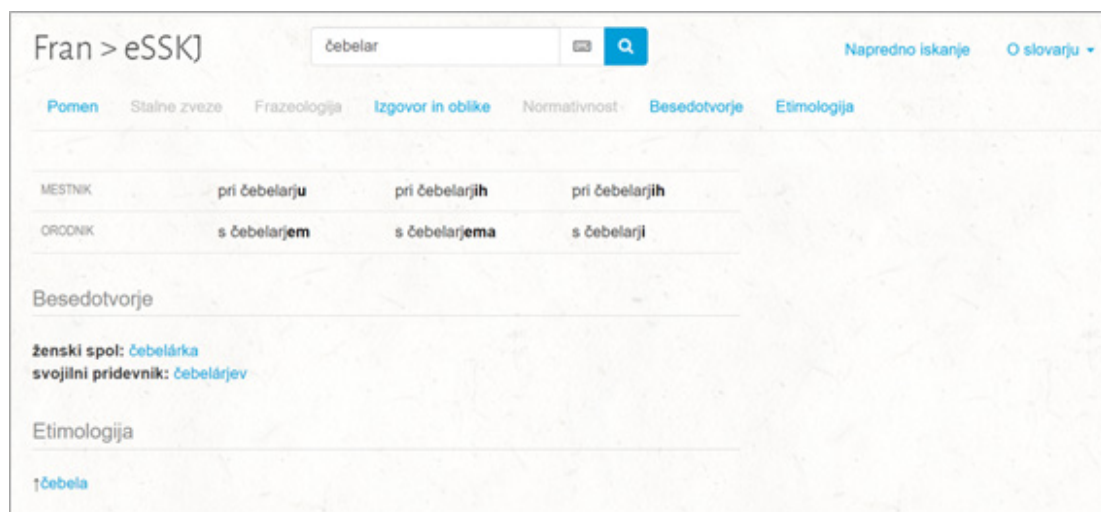


Fig. 1: Word formation tab for the entry *čebelar* ‘beekeeper’.

The word-formation tab also includes pairs such as *babica* ‘grandmother’ versus *dedek* ‘grandfather’. In these cases, the referencing is mutual; the female part of the

⁹ On the other hand, in the new Croatian dictionary *Hrvatski mrežni rječnik – Mrežnik*, the word-formation data are much more comprehensive. *Mrežnik* is a monolingual dictionary of standard Croatian, created from 2017 to 2021 as part of the research project of the same name, funded by the Croatian Science Foundation. Since 2022, the dictionary work has continued as an internal project of the Croatian Language Institute. The dictionary is available on the website www.rjecnik.hr/mreznik, and it includes three modules: a) a module for adult Croatian speakers, b) a module for elementary school students, and c) a module for speakers of Croatian as a second and foreign language. The module for adult Croatian speakers includes a section on word formation, which features the segmentation of a word into word-formation morphemes (e.g., *fizik-jar* for the entry *fizičar* ‘physicist’; phonetic changes are not indicated) and derivatives from the given entry (e.g., *fizičarev* ‘physicist’s’, *fizičarski* ‘physicist’s’, *fizičarka* ‘female physicist’). This is not limited to noun derivatives, as is the practice in the eSSKJ dictionary. Certain derivatives are also listed in the normative note tab, specifically when derivatives with different morphemes have different normative statuses (e.g., *čitalac* and *čitatelj* for ‘reader’). More about the principles of listing word-formation data can be found in the volume *Hrvatski mrežni rječnik – Mrežnik* or at <http://ihjj.hr/mreznik/uploads/5fcd02949bd03f2c37444dce7de2bfc.pdf> (Hudeček et al., 2016).

¹⁰ Both are direct derivatives of the verb *igrati* ‘to play’.

pair is listed in the word-formation tab of the masculine noun and vice versa. The same mutual connection exists for pairs for which the female derivative existed first due to social circumstances (e.g., *babica* ‘midwife’ vs. *babičar* ‘accoucheur’, *hostesa* ‘hostess’ vs. *hostesnik* ‘host’).

The lexicographer’s team decided to lower the frequency criteria for the inclusion of feminatives in the dictionary due to social responsibility. Feminatives (and masculatives) appear as independent entries when there is at least one occurrence in the corpus that can be used to create a dictionary example. For inclusion in the word-formation tab, one occurrence in the corpus is sufficient. The same frequency criteria apply for listing possessive adjectives and diminutives as independent entries as for other words. They are included in the word formation tab if they have at least twenty-five occurrences. Certain deviations are made for systemic alignment in cases of equivalent duplicates.

According to Concept II (Gliha Komac et al., 2015, p. 70), other primary derivatives from entries of different word classes can also be included in the dictionary if the analysis of the material shows that there are justified reasons for this (e.g., users’ communication difficulties, normative aspects, etc.). However, this practice was not established in the dictionary’s compilation. It is worth noting that the graded forms of adjectives, adverbs, and non-finite verb forms are systematically treated within inflectional-accentual patterns because they pertain to inflection rather than word formation. Data on secondary derivatives related to the entry are provided in the dictionary entries, where primary derivatives appear as headwords. Thus, the word formation tab shows at most one additional word-formation level. Exceptions are derivatives for which the intermediate formation stage exists only potentially but is not realized in the text; for example, *opičnjak* ‘ape-man’ will be listed under *opica* ‘ape’, not under the potential adjective **opični*.

Concept II also states that word-formation data, specifically the normative value of individual forms, are included in the normativeness tab (Gliha Komac et al., 2015, pp. 5–6). Examples include adjectival uses of nouns such as *aikido* (e.g., *aikido klub* ‘aikido club’), *kvir* (e.g., *kvir teorija* ‘queer theory’), and *reiki* (e.g., *reiki terapevt* ‘reiki therapist’), which are very common in contemporary material but are still understood by traditional normative studies as compounds rather than phrases. In the normativeness tab, there is a note: “The material shows that words like *video* are also used as adjectives (*video posnetek* ‘video recording’). Such examples are recognized in the 2001 Slovenian normative guide (§§ 487, 495, 498) as the first parts of compounds, and so writing them together (e.g., *videozapis* ‘video note’, *videoblog* ‘video blog’) is recommended, although separate writing is also allowed as non-preferred.” (More on this issue can be found in Gložančev, 2012, and Kern, 2012.)

In the future, word-formation data should definitely be included in the eSSKJ dictionary for other parts of speech, not just for nouns (similar to the way it is addressed in the Croatian Mrežnik dictionary).

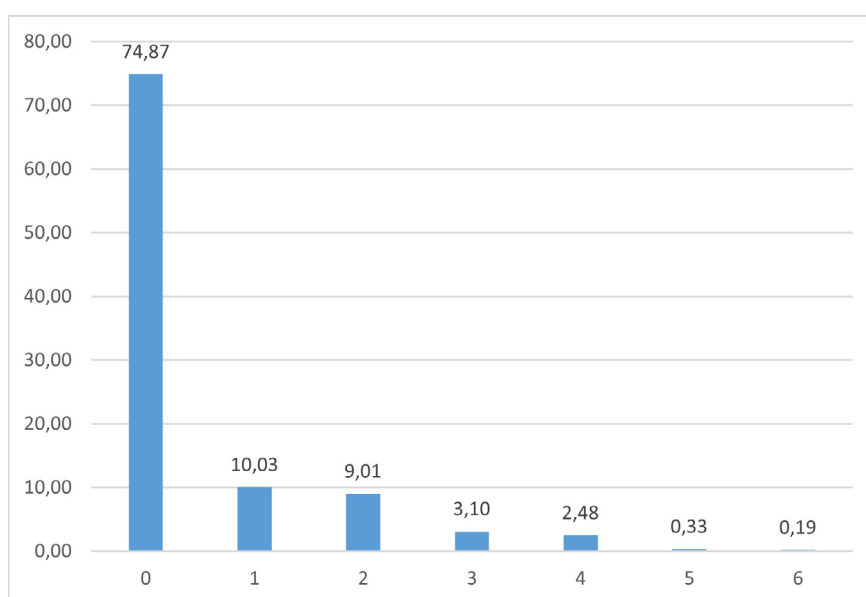
3. Occurrence of Words in Word-Formation Chains and Their Connotative Meanings

3.1 Occurrence of Words in Word-Formation Chains¹¹

As mentioned in the introduction, the multistage word-formation methodology seems to be particularly useful for analyzing material in all three mentioned phases. Based on digitized data from *Besednodružinski slovar slovenskega jezika: poskusni zvezek za iztočnice na B* (Slovenian Word-Family Dictionary: Trial Volume for Headwords Starting with B, 2004) by Irena Stramljič Breznik, which was prepared by Tomaž Erjavec,¹² a statistical analysis was conducted to determine how the word-formation stage affects the occurrence of words in the Gigafida 2.0 and metaFida corpora.

The analysis was initially conducted for all 11,136 derivatives in the word-family dictionary across both corpora. Subsequently, it was performed separately according to the part-of-speech (POS) classification of the simplex, specifically for noun, adjective, verb, and adverb simplexes (there are too few interjection and particle simplexes for the data to be statistically significant). This separation is justified by the fact that the POS classification of the simplex exhibits specific characteristics of the word-formation mechanism (cf. Skarżyński, 1999; 2003; Kern, 2017).

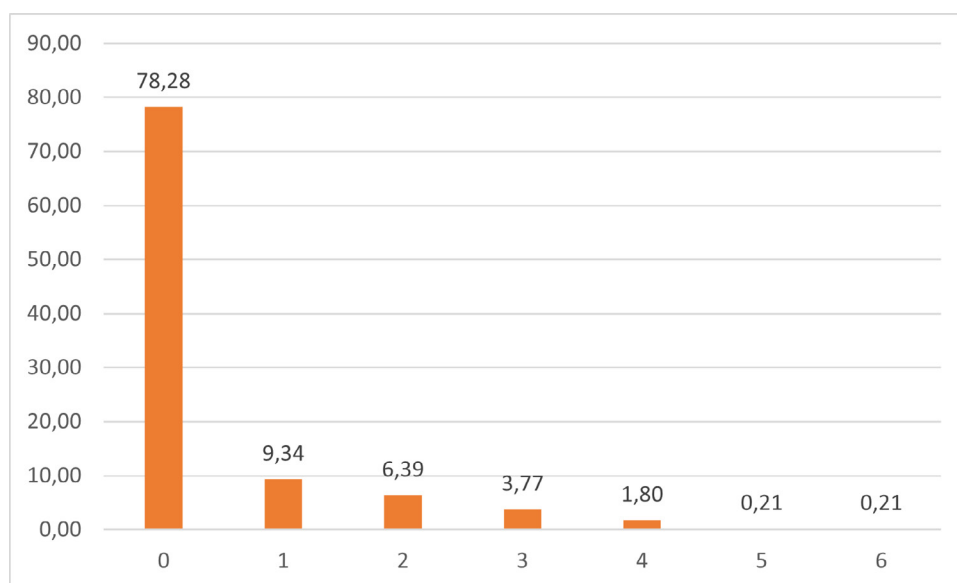
The analysis showed that the Gigafida 2.0 corpus (Graphs 1 and 3) and the metaFida corpus (Graphs 2 and 4) exhibit very similar data. In the *Slovenian Word-Family Dictionary*, the sixth word-formation stage is the highest, meaning that there is no examples of a word-formation chain with more than six stages. The data indicate that the frequency of first-stage derivatives relative to simplex words drops by a factor of 7.5. The occurrence of secondary derivatives relative to first-stage derivatives drops by one-tenth, and a larger drop occurs between the second and third word-formation stage, nearly threefold.



Graph 1: Occurrence of words (in %) according to the word-formation stage in the Gigafida 2.0 corpus

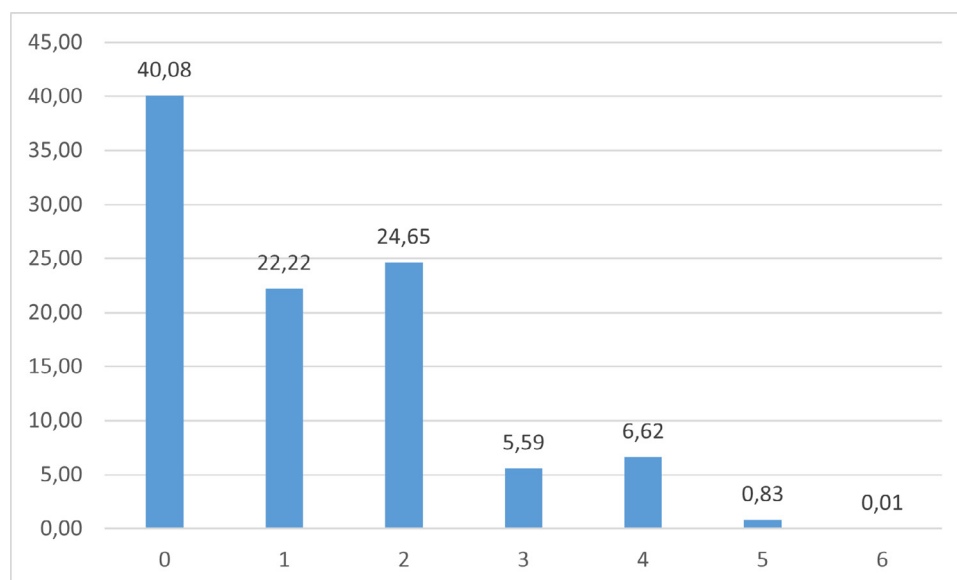
¹¹ The statistical data were prepared by Lazar Pavić.

¹² You can read more about the description of how the word-formation chains were detected in Erjavec et al., 2023.



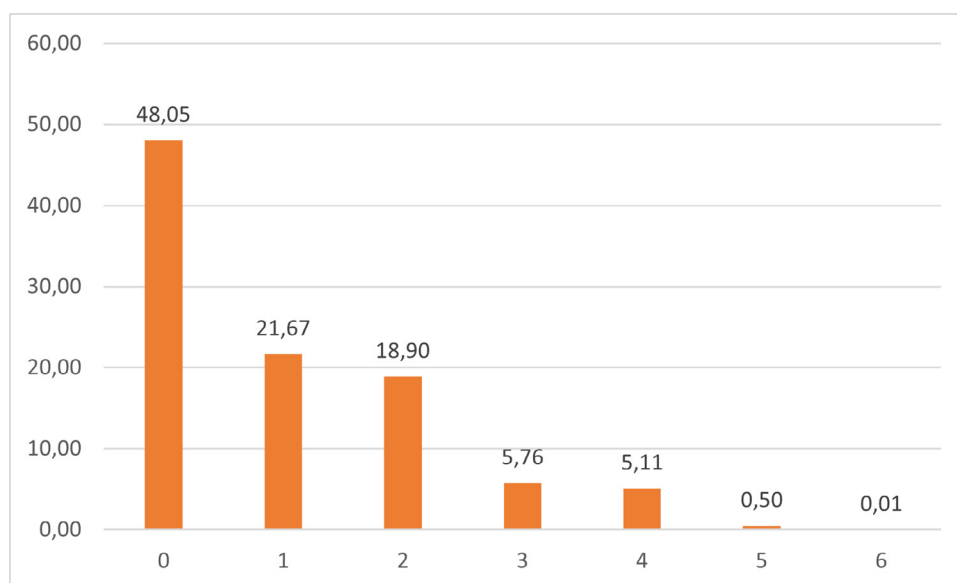
Graph 2: Occurrence of words (in %) according to the word-formation stage in the metaFida corpus

A similar pattern is also observed for nominal and adjectival multistage derivatives, although the data for verbal multistage derivatives differ slightly, and so they are highlighted separately (Graphs 3 and 4). In this group, the frequency of first-stage derivatives decreases by a factor of 1.8, but the frequency of second-stage derivatives, compared to first-stage derivatives, does not decrease but increases by one-tenth. However, a significant drop occurs between the second and third word-formation stages, specifically by a factor of 4.4, whereas the frequency of fourth-stage derivatives increases again.



Graph 3: Occurrence of verbs (in %) according to the word-formation stage in the Gigafida 2.0 corpus

The data from the metaFida corpus do not confirm the frequency increase of second- and fourth-stage derivatives. Instead, the frequency decreases with a similar exponent as in the graph for all words from the word-family dictionary.



Graph 4: Occurrence of verbs (in %) according to the word-formation stage in the metaFida corpus

The data provided are definitely informative and should be considered when setting criteria for including words in dictionaries. In particular, strategies need to be defined for handling higher-stage derivatives, specifically in terms of how much the fact that words belong to the same word family affects the reduction of the frequency criterion and to what extent this reduction is reasonable. It is not possible to provide universal guidelines for determining these thresholds for inclusion in the dictionary; even within a single dictionary, issues often need to be resolved on a case-by-case basis.

3.2 Connotative Meanings of Words in Word-Formation Chains

Previous research on connotative meaning within multistage derivatives (examining verbs of sensory perception; Kern 2018) has shown that both inherent and adherent connotative meanings occur among multistage derivatives. It has also been confirmed for Slovenian that the following depend on the (non)markedness of the word-formation base: 1) the size of the word-formation groups, 2) the number of word-formation models or the absence of certain types of word-formation models in stylistically marked word-formation bases, and 3) the length of word-formation chains. Word-formation groups with a neutral base are more branched than those with connotative meanings in Slovenian as well; word-formation groups with an expressive simplex motivate more multistage derivatives compared to all other bases. In word-formation families with connotative word-formation bases, connotativity is largely retained, with exceptions mainly in some primary verbal derivatives. Connotativity in word-formation groups in which the simplex does not have a connotative meaning primarily appears in first-stage compounds with interfixes (e. g., *drznogled* ‘bold-looking’, *krivogled* ‘cross-eyed’, *zvezdogled* ‘astrologer/astronomer’), (usually) second-stage nominal derivatives with zero suffixes (e. g., *dogled* ‘visibility’, *vpogled* ‘insight’, *oprez* ‘caution’), and higher-stage nominal derivatives (from the third stage onward) resulting from lexicalization (e. g., *pregledanec* ‘examinee’, *spogledljivec* ‘a flirtatious person’). There is less connotativity in word-formation

groups with nominal simplexes. It mainly appears at the second stage, mostly in verbs with prefixal suffixes, which are usually not productive for word formation. It is important to be consistent in dictionaries when providing dictionary qualifiers within word-formation chains. To achieve this goal, it is necessary to review the qualification of all derivatives in the word-formation chains.

4. Conclusions

In compiling dictionaries, considering word formation is essential for ensuring comprehensive and coherent entries. This study highlights the significance of multistage word formation in various phases of dictionary compilation, such as selecting criteria for including words, uniformity in dictionary definitions of related headwords, and consistency in marking stylistically significant senses.

The analysis conducted on the word-family dictionary and corpora, Gigafida 2.0 and metaFida, demonstrates the impact of the word-formation stage on word frequency, which is very informative in compiling dictionaries. The findings indicate a significant decrease in the frequency of derivatives as the formation stage increases, with some variations observed across verbal multistage derivatives. Notably, whereas noun and adjective derivatives follow a consistent pattern, verb derivatives exhibit unique trends, particularly in the second and fourth stages of formation. The data provided should be considered when setting criteria for including words in dictionaries. In particular, strategies need to be defined for handling higher-stage derivatives, specifically in terms of how much the fact that words belong to the same word family affects the reduction of the frequency criterion and to what extent this reduction is reasonable.

Furthermore, previous research on the connotativity of multistage derivatives confirms the presence of both inherent and adherent connotative meanings. The markedness of the word-formation base influences various aspects, such as the size of word-formation groups, the number of formation models, and the length of formation chains. It is important to be consistent in dictionaries when providing dictionary qualifiers within word-formation chains.

In conclusion, integrating word-formation data into dictionary entries enhances the lexicographic work, providing users with a richer and more accurate linguistic resource. Future research should continue to explore the nuances of word formation in Slovenian and other Slavic languages, contributing to the development of innovative and comprehensive dictionaries.

Acknowledgements

This article was written as part of the project Formant Combinatorics in Slovenian (J6-3134), funded by the Slovenian Research and Innovation Agency (ARIS). ARIS funding was also provided through the core program The Slovenian Language in Synchronic and Diachronic Development (P6-0038).

References

- August, G. (1993). New Trends in the Research on Word-Family Dictionaries. *Studia Anglica Posnaniensis*, 25–27, 183–197.
- Burkacka, I. (2012). *Kombinatoryka sufiksalna w polskiej derywacji odrzeczownikowej*. Warsaw: Wydział Polonistyki Uniwersytetu Warszawskiego.
- Čibej, J. (2021). Strojno luščenje medbesednih povezav v oblikoslovnem leksikonu Sloleks 2.0. In Š. Arhar Holdt (Ed.), *Nova slovnica sodobne standardne slovenščine: viri in metode* (1st ed., pp. 125–159). Znanstvena založba Filozofske fakultete. <https://e-knjige.ff.uni-lj.si/znanstvena-zalozba/catalog/view/325/477/7313-1>.
- Dobrovoljc, K., Krek, S., & Erjavec, T. (2017). The Sloleks Morphological Lexicon and its Future Development. In *Dictionary of Modern Slovene: Problems and Solutions*. Znanstvena založba Filozofske fakultete. <https://doi.org/10.4312/9789612379131>.
- Erjavec, T., Pranjić, M., Pelicon, A., Kern, B., Stramljič Breznik, I., & Pollak, S. (2023). Automating derivational morphology for Slovenian. In M. Medved et al. (Ed.), *eLex 2023 : electronic lexicography in the 21st century (eLex 2023) : proceedings of the eLex 2023 conference: [Brno], 27–29 June 2023. Brno: Lexical Computing CZ, 2023* (pp. 449–465). <https://elex.link/elex2023/wp-content/uploads/92.pdf>
- eSSKJ. (2017). Predgovor. https://fran.si/201/esskj-slovar-slovenskega-knjiznega-jezika/datoteke/eSSKJ_Predgovor.pdf.
- Gigafida 2.0: Korpus pisne standardne slovenščine*. virij.cjvt.si/gigafida.
- Gliha Komac, N., et al. (2015). *Koncept novega razlagalnega slovarja slovenskega knjižnega jezika* [electronic source]. Založba ZRC, Znanstvenoraziskovalni center SAZU.
- Gložančev, A. (2012). Novejša slovenska leksika v luči obravnave samostalniških zloženek v Slovenskem pravopisu 2001. In N. Jakop, & H. Dobrovoljc (Eds.), *Pravopisna stikanja* (pp. 125–150). Založba ZRC, Znanstvenoraziskovalni center SAZU.
- Hudeček, L., & Mihaljević, M. (2016). Tvorbeni podatci. In L. Hudeček, & M. Mihaljević (Eds.), *Hrvatski mrežni rječnik – Mrežnik*. Institut za hrvatski jezik i jezikoslovlje. <http://ihjj.hr/mreznik/uploads/06c8a451ca3cf1d9114ad3744afb4a3e.pdf>
- Kern, B. (2010). Stopenjsko besedotvorje. *Slavistična revija*, 58(3), 335–348.
- Kern, B. (2012). Pisanje skupaj in narazen v Slovarju novejšega besedja slovenskega jezika. In N. Jakop & H. Dobrovoljc (Eds.), *Pravopisna stikanja* (pp. 141–151). Založba ZRC, Znanstvenoraziskovalni center SAZU.
- Kern, B. (2017). *Stopenjsko besedotvorje (Na primeru glagolov čutnega zaznavanja)*. Založba ZRC.
- Kern, B. (2018). Konotativnost z vidika stopenjskega besedotvorja. *Slavistična revija*, 66(2), 130–141.

Krek, S., Kosem, I., & Gantar, P. (2013). *Predlog za izdelavo Slovarja sodobnega slovenskega jezika* [electronic source]. http://www.sssj.si/datoteke/Predlog_SSSJ_v1.1.pdf.

metaFida v1.0. https://www.clarin.si/noske/run.cgi/corp_info?corpname=mfida10&struct_attr_stats=1

Olejniczak, M. (2003). 30 lat słotwórstwa gniazdowego (rys historyczny). In M. Skarżyński (Ed.), *Słotwórstwo gniazdowe: Historia, metoda, zastosowania* (pp. 12–28). Księgarnia Akademicka.

Skarżyński, M. (1999). *Powstanie i rozwój polskiego słotwórstwa opisowego*. Instytut Filologii Polskiej Uniwersytetu Jagiellońskiego.

Skarżyński, M. (2003). Czasowniki jako centra gniazd słotwórczych. In M. Skarżyński (Ed.), *Słotwórstwo gniazdowe: Historia – metoda – zastosowania* (pp. 41–55). Księgarnia Akademicka.

Stramljič Breznik, I. (2004). *Besednodružinski slovar slovenskega jezika: Poskusni zvezek za iztočnice na B*. Slavistično društvo.

Vidovič Muha, A. (2018). *Slovensko skladijsko besedotvorje*. Znanstvena založba Filozofske fakultete.

Contact information

Boris Kern

ZRC SAZU, Fran Ramovš Institute of the Slovenian Language

University of Nova Gorica, School of Humanities

boris.kern@zrc-sazu.si

