



EURALEX XIX

**Congress of the
European Association
for Lexicography**

Lexicography for inclusion

7-9 September 2021
Virtual

www.euralex2020.gr

**Proceedings Book
Volume 2**

Edited by Zoe Gavriilidou, Lydia Mitits, Spyros Kiosses

EURALEX Proceedings

ISSN 2521-7100

ISBN 978-618-85138-2-2

Published by: SynMorPhoSe Lab, Democritus University of Thrace

Komotini, Greece, 69100

e-edition

Publication is free of charge

Edited by: Zoe Gavriilidou, Lydia Mitits, Spyros Kiosses

English Language Proofreading: Lydia Mitits and Spyridon Kiosses

Technical Editor: Kyriakos Zagliveris



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License

2021 Edition

New words in old sources: Additions to the lemma list of a historical scholarly dictionary

Johannsson E.T., Battista S.

Department of Nordic Studies and Linguistics, University of Copenhagen
ellert@hum.ku.dk, sb@hum.ku.dk

Abstract

This paper accounts for recent additions to the lemma list of *A Dictionary of Old Norse Prose (ONP)*, which is a historical dictionary describing the medieval language of Iceland and Norway. The dictionary was established in 1939 and has throughout the years built up a large database containing about 800.000 example citations illustrating the vocabulary of all prose genres. The lemma list consists of about 65000 words with accompanying citations, but is continuously being revised. After giving a brief account of the history of this project we give an overview of the editorial principles, the criteria used for defining a lemma and discuss different types of lemmas found in the dictionary. We describe the characteristics of entries in ONP and mention different types of entries found in the online version. We then focus on the period from 2010-2019 and present a study into new additions to the lemma list during those years. We analyze these more recent words, divide them into eight groups and give some examples that illustrate the processes involved when new headwords are established. The results of the study show that most of the later additions to the lemma list come about in relation to editorial work on other words. A significant proportion of new words are established when new compounds are identified while editing un-compounded, simplex words, but other factors are in play as well.

Keywords: historical lexicography; morphology; lexicology

1 Introduction

A Dictionary of Old Norse Prose (ONP) is a dictionary project hosted at the University of Copenhagen and part of the Arnamagnæan Institute of Old Norse Manuscript Studies. This dictionary accounts for the vocabulary of the language of medieval Iceland and Norway, from around 1150 to 1370 (Norway) and to 1540 (Iceland). The corpus consists of texts preserved in manuscripts, with all the implications of text transmission, which make every version of a text unique. The lexical material is the result of extensive excerption work mostly from scholarly editions of manuscript texts and in some cases directly from manuscripts. ONP has since 2010 been available as an online resource at onp.ku.dk, which provides access to the material from the published volumes of the dictionary (1995-2004) as well as more recently edited dictionary entries and unedited dictionary material. The work on the dictionary continues with new entries published online, as well as addition of new features to the online version.

The paper is organized as follows: After giving a brief account of the background and history of the project we discuss some of the editorial principles, the criteria used for defining a lemma and account for different types of lemmas in ONP. Next, we describe the characteristics of entries in the dictionary and give examples of the different types of entries found in the online version. We then account for a study into new additions to the lemma list of ONP during the period from 2010 to 2019 and present the results of our findings.

2 Background

The ONP dictionary project was established in 1939. The focus of the lexicographic work has always been on the language of medieval prose texts from Norway and Iceland as the poetic language had already been described in great detail with the publication of a revised dictionary of the poetic language a few years earlier (Jónsson 1931). The dictionary was originally meant to be a supplementary to the *Ordbog over det gamle norske Sprog* (1886-1896), by Johan Fritzner. In the meantime, it had become clear that Fritzner's work had some limitations with its normalization practices and citing of many text editions that had become obsolete. It was decided that a new lexical description of Old Norse prose was needed, with the aim of giving an exhaustive representation of the vocabulary of Old Norse prose excerpted from all known texts in a scholarly edition or directly from manuscripts. A new dictionary of this kind would also fit well into the Danish lexicographic tradition and would strengthen Copenhagen as an important research center for the Nordic cultural heritage, with many important manuscripts being preserved there.

In the first decades, the dictionary staff was mostly concerned with gathering material for an eventual print publication. This meant selectively excerpting all known medieval texts, representing the vocabulary of different Old Norse prose genres, by collecting examples of word use. Selected citations were written onto slips, which were then filed under a particular headword in alphabetical order (cf. figure 1). The citation collection was intended to be very detailed and illustrate the range in meaning of every word.

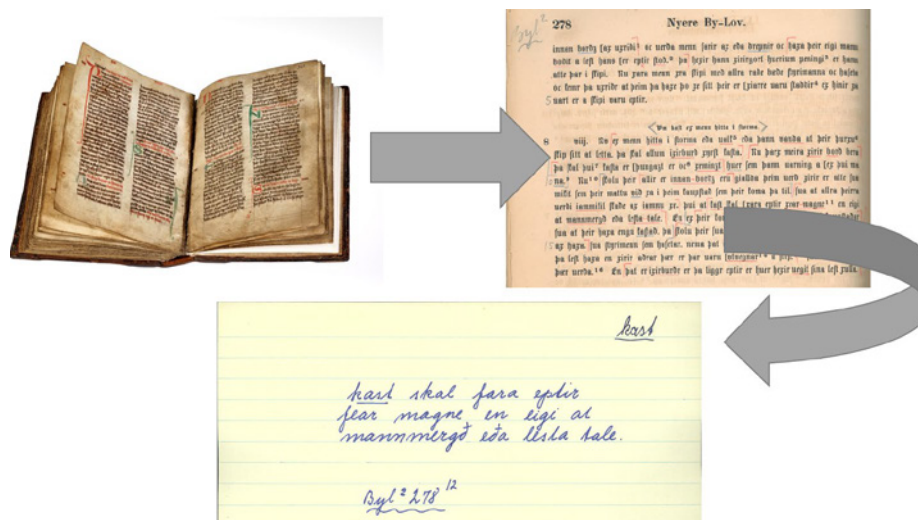


Figure 1: The medieval manuscript is edited in a scholarly edition, which in turn is excerpted by underlining relevant citations and writing them down on a paper slip.

A few key works – representing different genres, from particularly relevant manuscripts – were exhaustively excerpted, i.e., each and every word in those texts was copied onto a slip along with its syntactic context and filed in the dictionary archive. This provided additional examples of many commonly used words that were underrepresented in the citation collection. The exhaustively excerpted texts represent different genres and among these we find a section of *Snorra Edda* (mythological tales) from the manuscript GKS 2367 4^o, a fragment of *Egils saga Skallagrímssonar* (an Icelandic family saga) from AM 162 A 0 fol, and *Íslendingabók* (an historical account of the settlement of Iceland) from AM 113 fol^x.¹ As the decades passed and more scholarly text editions were published, the citation collection grew. This eventually resulted in an archive consisting of around 750.000 handwritten slips, organized under 65.000 lemmas (cf. Johannsson & Battista 2014).

Once the examples had been collected, plans were made for the publication of a multi-volume print dictionary. The first volume was an index volume published in 1989 followed by three volumes of dictionary entries covering the alphabet from *a-em*. In 2005, after the third volume had come out, the print publication was put on hold and preparation began for the transition to a digital online dictionary (cf. Johannsson & Battista 2016). ONP Online (onp.ku.dk) was launched in 2010 and combined material from the already published volumes, newly edited entries and unpublished material in the form of scanned citation slips (cf. Johannsson 2019). The online version has more recently been redesigned and enhanced in different ways with linking to other digital resources for Old Norse (cf. Wills & Johannsson 2019).

The editing work is ongoing and focuses on word groups rather than alphabetical order of lemmas. The headwords are divided into twelve groups for the purpose of editing: simplex nouns (with fewer than ten citations) simplex nouns (with ten or more citations), compound nouns, verbs, simplex adjectives, compound adjectives, simplex adverbs, compound adverbs, pronouns, numerals, conjunctions and prepositions (cf. Johannsson and Battista 2016). The grouping is mostly based on part of speech, but also morphological features and frequency. The first group to be edited following this new procedure was simplex nouns with more than ten citations, followed by simplex nouns with fewer citations. Since this new editing procedure was put in place all the simplex nouns have been edited along with simplex adjectives and adverbs. Pronouns, numerals, conjunctions and prepositions are also close to being finished. The editing of verbs is done in two rounds with an initial round of editing focusing on argument structure and formal categories rather than meaning. This initial editing is now completed. The second round, which involves the semantic editing of verbs, is currently underway. The largest groups that remain completely unedited are compound nouns and compound adjectives.

3 Editorial Principles and Lemma Criteria

The ONP dictionary has from the beginning followed certain clearly defined editorial principles. An important feature, which distinguishes ONP from its predecessors, involves adhering to the original orthography of the source texts and maintaining rigorous philological standards. This entails that the citations are as far as possible taken from diplomatic scholarly editions or even unpublished manuscripts, which means that the orthography of the citation examples is highly irregular (cf. Johannsson & Battista 2016: 118-119).

Even though the example citations are not normalized, the lemma list of ONP is normalized according to a normalization standard developed by ONP. This normalization is similar to the classic Old Norse spelling often used in text editions, which reflects the phonological state of Icelandic around year 1200 (a detailed overview of normalization practices and the principles of Old Norse normalization is found in Bernarðsson et al. 2019). ONP's orthography differs in some significant ways as it tries to take into account both Norwegian and Icelandic language development. It does not reflect some special Icelandic sound changes, such as vowel lengthening before certain consonant clusters, e.g., standard Old

¹ See, e.g., handrit.is for more information about these manuscripts and shelf marks.

Norse *úlfr* ‘wolf’ with a long u-vowel is normalized as *ulfr* by ONP as this lengthening rule did not take hold in Norway. Another important difference is the consistent use of the acute accent to mark all long vowels by also using the lesser-known characters *é* and *ó*, which are not part of the traditional Old Norse standard orthography. This approach has some pedagogical advantages and constitutes a good compromise when it comes to developing standards that can be used for editions of both Norwegian and Icelandic texts. ONP’s orthography, however, has only a limited tradition in text editions. It has also been revised a few times, most recently in 2003, which could give the impression that it is not as well established as other orthographic standards (cf. discussion in Johannsson and Battista 2020).

When choosing the form of lemma, ONP follows the criteria already established by its predecessors. Nouns are listed in their nominative singular form, verbs are given in the infinitive, active voice and adjectives in the masculine, nominative singular form. (cf. ONP Nøgle/Keys 2004)

The main components of a typical dictionary entry is the headword, along with grammatical classification, details of inflection and the example citations found in the dictionary archives. The citations are either typed in or shown as scanned paper slips (see section 4 below for detailed look at different types of entries). Unlike most dictionaries, ONP lists all the examples it has registered for each lemma, and, in many cases, these are all the attested examples of word use.

The ONP dictionary features several different types of entries and this is reflected in the lemma list. There is a distinction between so-called standard headwords and secondary headwords. Standard headwords contain citations arranged chronologically according to senses (if the entry has already been edited), and a concluding section with supplementary information. The secondary headwords are in principle “registrations of a word’s existence, with no semantic explanation, but with references to other dictionaries and glossaries (Gloss.), and occasionally to secondary literature” (cf. ONP Nøgle/Keys 2004). These secondary entries can be further divided into several subtypes:

- *Poetical words.* These are words, which, in spite of appearing only in poetry context, are recorded in the dictionary for their lexicographic value. These entries only rarely have example citations (from poetic use in prose texts). They contain reference to a relevant glossary over the poetic language and are labelled (*poet.*). An example would be *oglir*; a word for hawk or falcon with multiple occurrences in poetry but never in prose.
- *Non-assimilated foreign words.* These are foreign words that appear in an Old Norse context, but are not adapted to the language, either phonologically or morphologically. Such words are labelled (*alien.*) or (*foreign*) and spelled according to the language of origin. Some examples would be *cherub*, *schismus* and *synecdoche*. Integrated loanwords however are treated in the same way as other Old Norse words.
- *Starred words.* These words are of various sorts. They can be words that appear in other Old Norse dictionaries, but fall outside ONP’s defined scope. They can also be so-called “ghost words” which are the result of an erroneous reconstruction in a text edition or a misreading of a manuscript, which has found its way into a published text, or are based on an interpretation with which ONP does not agree. Many such words stem from earlier dictionaries, based on material from old and obsolete editions that have since been replaced by more precise scholarly editions. A good example is the hapax *duma*, which is a result of a misreading of the quite common verb *dvína* in an early text edition. It is often difficult to see the difference between *-in-* and *-m-* in manuscripts so such errors can arise. A more careful reading in a later edition has revealed the mistake. The word *duma* is still recorded but receives a star, as it does not have any philological basis.
- *Questionable words.* Some words of uncertain status, which are attested in the actual medieval material, but are most likely a result of an error, are preceded by a question mark and are usually followed by a reference to a likely “correct” form. An example would be the adjective *?gozkr* for *girzkr* ‘Greek’
- *Prefixes and suffixes.* These are items listed in other dictionaries/glossaries. An example is the suffix *-geðjaðr*, which only exists in derived words such as *lausgeðjaðr* ‘indecisive’.
- *References.* There are references from so-called alternative forms, which are variant forms that share most characteristics with the main form, i.e., high frequency and straight forward normalization, as well as references to so-called special forms, which are isolated occurrences that are not suited to normalization but cannot be a result of an error. An example would be *skemmtan* which is a common alternative form to *skemmtun* ‘entertainment’, which is linked to a single entry *skemmtun*, *skemmtan sb. f.*
- *Final element of a compound.* Such elements are not attested as independent words and thus have no definition or citations, but only a reference to the attested compounds where the element occurs, as well as relevant references to older dictionaries and secondary literature. An example would be *-saltaðr* ‘salted’, which is only attested in the adjectival compound *ósaltaðr* ‘unsalted’.
- *Out of scope words.* These are words that have been registered by ONP but do not fall within the defined scope of the dictionary. These words can be for example words that are only attested in younger text sources but are deemed to have relevance to the description of the medieval vocabulary, such as *allraheilagramannamesuaftann* ‘the eve of all saints mass’ only attested in a Norwegian text from around 1380 but shows an example of a very long nominal compound. Another type would be place names and personal names, such as *England* or *Egill*. The words that fall under this heading are clearly marked as such.

The standard headwords follow certain principles. As stated in the User’s Guide for the dictionary the oldest and etymologically most original form is usually chosen as the main form.² If a variation between forms is purely

² The User’s Guide is a helpful aid originally intended for the users of the print dictionary. It explains the editorial principles of ONP in some detail as well as the structure of the entries. The guide was published in a booklet that accompanied each printed volume, and contained corrections along with a list of abbreviations and symbols. The latest one is ONP Nøgle/Keys 2004. The User’s Guide is now

orthographical or a result of a clearly understood phonetic variation, i.e., dialectal, the variants are treated under the same headword. If the difference between variant words is the result of inflectional discrepancies, the words in question will each be listed in their own entry with their own set of citations. This is often the case with verbs that are attested with forms belonging to different conjugational patterns or nouns that display forms that belong to different declensional classes. This is also the case with compound words where the members of the compound are the same but joined together by different morphological elements, e.g., *barns-faðir*, *barna-faðir* and *barn-faðir*, which all have the same basic meaning as ‘a father of child/ren’. There are two exceptions, where variant forms are grouped together under the same lemma. One is words where the second element starts with an *s* and variant forms with both *-s-* and *-ss-* are recorded the lemma is normalized as with one *s* in brackets, e.g., *dóm(s)æti* ‘court seat’. Another rare exception is when the first member ends in *-ar* and alternatively in *-a* in some Norwegian sources, e.g., *atfara(r)þing* ‘assembly called to obtain an order of distraint’ (cf. ONP Nøgle/Keys 2004).

4 Types of Entries in ONP Online

In its current form, the online dictionary displays one of four types of entries for each standard headword. The reason for these different entry types can be explained by the complicated history of the project as a partial print publication and later on as an online work in progress. In the sections below we will give a brief description and examples of each type.

4.1 Edited Entries from the Print Edition

The edited entries that also appear in the printed volumes are published online without any major changes. The structure contains all the same detailed information found in print as well as the keyed-in citation examples. Unlike the print edition all the citations that accompany each headword are listed and most citations are displayed along with a scanned page from an edition of the text (cf. figure 2). When publishing in print the number of citations often had to be reduced so the editors would pick the best representative examples for publication. In the online version of ONP all the citations found in the dictionary archive have been made available and the unpublished citations have been fully integrated into the entry structure and placed under the appropriate sense.

The figure consists of two overlapping screenshots of the ONP online dictionary interface. The background screenshot shows the entry for 'ari sb. m. [-a; -ar]'. It includes a search filter for 22 citations, a list of citations, and a section for '2) (om afbildning/skugga af of a depiction/shadow)'. A red box highlights a citation example: 'AM 619 4* (5-1200-1220) 44v(33)-45v(30) = Hinnilar : Hinnilar & ættmenn'. The superimposed screenshot shows a scanned page from the print edition of the entry, with a red box highlighting the same citation example: 'AM 619 4* (5-1200-1220) 44v(33)-45v(30) = Hinnilar : Hinnilar & ættmenn'. The scanned page also includes a table of references at the bottom.

Figure 2: Screenshots showing an entry that was previously published in print. Additional citations have been added to the entry structure and for each citation a scanned page from the relevant edition is made available (second screenshot superimposed).

4.2 Entries that Have Been Edited since the Print Edition Was Put on Hold

The entries that have been edited in the period since the print edition are very similar to the ones described above. All the available citations have been organized and placed in the entry structure. For many headwords selected citations have been marked with three bullets (●●●). These show citations that the editor of the relevant entry has highlighted to demonstrate a particular representative usage of the word (cf. figure 3). In addition to the keyed-in citation examples, the scanned slips are also visible and the extra information they may contain is accessible in this way.

ONP: Dictionary of Old Norse Prose

Home Words Info Indices

Home Words Info Indices Manuscripts Works Bibliography

kast sb. n. [k.ˈsɑt]

var. *Gisl 37^a: NKS 1181 fol^a "S"*

NKS 1181 fol^a (477a-104) 53v-70r — *Gísla saga Sölvunar*

(Finna Jónsson 1929: 1-75)

1) *kast* (m. stein, etc.) // stone-throwing

2) *overholdastning* (af lading) // throw

3) *det at lagge fiskerast ud* // casting of a fishing net

4) *terningkast* // throw of dice

Utm *kast*

Gísl tekr upp stein einn ok kastar ut í holm

••• Sock i þu skipti ríðu til geyra ok rak s

kast skál fara eptir fear magna en eigi at má

••• vm kast a skípi of ofháli et

kast skál fara eptir fjármaga en eigi at má

þu eignast þeir sild allu ok kast [var. varp *Höfn þing* 32 4^a "Gr", etc.] sm

••• Ein enn þeir kútu er þu skal sva flyta ex varst sem fundast sáham þat er ... þorkoner ok þreita ok kast [var. cast *Kp* 5^a] ex viðloger þat af þessum ... *Kp* Page 130^a

Land 137^a

Figure 3: Screenshots showing an entry that was edited for ONP Online. Three bullets highlight selected citations and each citation is provided with a scanned citation slip as well as a scanned page from the relevant edition (second screenshot superimposed).

4.3 Structurally Edited Verb Entries

After the print edition was put on hold, it was decided that the editing of verbs should be done in two stages. The first stage would include structuring the verbs, i.e., grouping the citations according to structural criteria, e.g., participles, active and middle voice forms as well as verbal clitics. The entry consists of detailed information about the verbal argument structure and the accompanying citations are displayed as scanned slips. However, there is no information about meaning and the citations have not been keyed in (although sometimes the actual form has been entered (cf. figure 4)). All the verbs have been structurally edited in such a manner. The second stage of the editing of the verbs has recently commenced and includes semantic structuring of the verbs in line with what is found for verbs in the print edition, but such fully edited verbs have not yet been published online.

ONP: Dictionary of Old Norse Prose

Home Words Info Indices Manuscripts

Home Words Info Indices Manuscripts Works Bibliography

skara vb.

Ehrlajf [2003] 299^a

WaldAug 9 10 4^a 4100 170 1-24r — *Egloga saga*

(New 2003) [RA A 15], 85-121; *lagræðing* (S 3-149) *ovevst* (Ska, 7151-225 ...)

[e-mte-u] [t-á-e-t] : [e-t] [mte-e-jum]

skaraelk

[var. skara (var skara NKS 299^a) varðtræði með þem skardu frá skardu Látú: 219]

• skara eld

(see citation slip for text)

• 1 *impers: er skarar* (með) skjöldum [fyrir 1]

(see citation slip for text)

(see citation slip for text)

• 1 *part. part: skaradr*

horða voru skaradr ok bátt með bráðk

• skaradr skjöldum

(see citation slip for text)

(see citation slip for text)

*my bh kast valdgan ok vel tvítt [var. skipti með stafa skarlagan skjöldum ... GKS 109 fol^a P^a; skjöldum daga með stafa A... var *F46F 150*; AM 142 fol^a R^a]

B. skara + *prep./adv.*

Figure 4: Screenshots showing a verb entry that has been partially edited to illustrate structural features. There are no definitions and most citations have not been keyed in. Each citation is displayed as a scanned citation slip accompanied by a scanned page from the relevant edition (second screenshot superimposed).

4.4 Unedited Entries

In the first version of the online dictionary, the majority of entries was of this kind, consisting only of headword along with the accompanying scanned slips with no entry structure and only a list of the citations in chronological order (cf. figure 5). As the editing work has progressed this entry type is most common for compound nouns and adjectives, which are the largest groups of words that have not yet been edited. Eventually, as the editing work progresses, these types of entries will gradually be replaced by fully edited and structured entries.

ONP: Dictionary of Old Norse Prose

Home Words Info Indices Manuscripts Works Bibliography

steinkast sb. n.

Accepted Citations slip

Citations by ms. date Citations by source

Search/filter 6 citations...

wa læng i dǫpnum rennaðs sem fœgr steinkast of 'af
hann kom nær steinkast frá heilnum
þá grek hann frá postulum svá sem steinkast læng
(see citation slip for text)
(see citation slip for text)
(see citation slip for text)

cf. simplex kast (27)

Form: (N); steinkast (1); steinkast (1); steinkast (1);
Gloss: E; CD; CW Suppl; Suppl; Fr; NO; BI

Genre (expected): religious works (inf: 4 (2)); family sagas (inf: 1 (1)); legendary sagas (inf: 1 (0)); learned works (inf: 0 (0)); legal works (inf: 0 (1)); contemporary sagas (inf: 0 (0));
unclassified (inf: 0 (0)); romances (inf: 0 (1)); þættir (inf: 0 (0)); charters (inf: 0 (0)); historical works (inf: 0 (1));

Show timeline of citations by manuscript date and genre

Word in other corpora Framer Cleanby/Zoega Bionda

Figure 5: Screenshots showing an entry that has not been edited. There is no entry structure and most citations have not been keyed in. The citations appear in chronological order. Each citation is displayed as a scanned citation slip accompanied by a scanned page from the relevant edition (second screenshot superimposed).

4.5 The Ideal Entry

Only the first two types of entries discussed above (4.1 and 4.2) are representative of the ideal form of all entries once the editing work has been completed. An ideal entry in the ONP should include the following features:

- The headword in a cardinal form in normalized orthography.
- Morphological information (inflectional pattern and verb conjugation) based on texts (mainly the actual examples found in the dictionary database).
- Semantic tree.
- Two target languages: Danish and English.³
- References to foreign parallel texts (esp. Latin).
- Keyed-in citations with non-normalized orthography, i.e., the orthography of the relevant manuscript or scholarly edition is rendered as closely as possible, with frequent use of special characters.
- Detailed system of sigla indicating not only reference to an edition but also the actual manuscript for each section of the text (in some cases different manuscripts are used within the same edition).
- Syntactic information (especially verb complements and prepositional use).
- Phrases and collocations.
- References to glossaries and, where relevant, references to secondary literature.
- Scanned editions and/or links to images of manuscripts.

Currently, about half of the entries in ONP Online have all or most of the features listed above. For the other half, the editing work is ongoing and only several of the features are already available. In addition to the features listed here ONP Online provides the users with various innovative ways to access the data from the dictionary (cf. e.g., Wills & Johannsson 2019) as well as links to various secondary sources, such as earlier dictionaries, electronic text editions and scanned manuscript images.

5 The Current Study

Since the textual material is known and clearly defined in space and time, we would expect the list of lemmas to be stable as there is in principle no new material being added to the text corpus. However, it turns out that the lemma list of the dictionary continues to evolve with some words being removed as well as new words being added. In its current form the lemma list consists of about 90.000 items and the database contains about 800.000 example citations. About 65.000 of these items are associated with citations, the other items on the lemma list are various types of secondary headwords, which do not contain any examples of usage.

In order to better understand the dynamics of the lemma list and how new items are added to it we conducted a small study. We decided to limit the study to a ten-year period from 2010-2019. As all the lemmas in ONP are organized in a database and each item in the lemma table is assigned a number, it was relatively easy to figure out the additions for the defined time period. After we had filtered out obvious mistakes and entries that were old but had received a new number, we were left with 3789 new items. Most of those items turned out to be secondary headwords of the type that do not contain any citations. Of those there were 2855 references to alternative forms or side forms as well as 194 items, which were suffixes, affixes and second members of compounds. Additional 153 items were different types of secondary entries without citations.

³ The entries that also appeared in the print edition have two target languages. The entries that only have been published online are still mostly monolingual (either in Danish or English), but work is ongoing to add the other target language where it is missing.

We were then left with 586 new words that had some citations associated with them. Not all of these could be classified as standard headwords as some secondary headwords also have citations. We found that out of the 586 words with citations 42 were outside the scope of the dictionary. Further 100 were labeled questionable and we decided not to account for them further, as these headwords are inherently problematic. The remaining 444 words are standard headwords and can be divided into eight groups:

Compounds added from examples of simplex words (208 examples)

This is the largest group of words, as could be expected, since in Old Norse, as in all Germanic languages, compounds are a very productive word category. Besides the morphological variation in the manuscript material variable factors are whether a compound has been written as one or two words and whether phonological changes have occurred at the juncture of the two elements of a compound. These variables can reflect individual scribal traditions but also the degree of lexicalization of a given compound (cf. Bakken 1995: 170 ff.), or in other words the extent to which it was considered a semantic unit. An example of newly added compound is the hapax *þurfandahjólþ* ‘help for the needy’, a nominal compound the first element of which is a noun in the gen. pl. This word has been added to the lemma list while editing the simplex *hjólþ* ‘help’, as it has been considered a neologism created on the basis of the foreign Latin model *Auxilium Egentium*. There are also examples of adjectival compounds such as *nýgerðr* ‘newly done’, and *nýklyppðr* ‘newly shaved’, which are formed by the adjective *nýr* ‘new’ + participle. We also find examples of compound verbs, which are often prepositional verbs with occurrences where the conjugated form has a preposition as the first element, for instance *upplíða* vs. *líða upp*, *upptelja* vs. *telja upp*, or *viðkennask* vs. *kennask við*.

Homograph reorganization (43 examples)

Another significant group of new words is the result of homograph reorganization. All in all, we found 42 examples of a new word that had been added to the lemma list where an identically spelled counterpart already existed. In most cases, this is a result of an editing process of a particular headword where either morphological or semantic evidence has suggested that the examples should be divided up between different homonyms that fulfill the criteria for an independent headword. There are many examples here of verb forms that show the characteristics of a different conjugational pattern than a more established type. An example would be the verb *líka* ‘close’ which usually is a strong verb with a vowel change in the root, present form *lýkr* ‘closes’, but ONP has recorded one example with a present form *líkar*, which indicates a weak verb conjugation and has therefore given rise to a “new” weak verb *líka*². Slightly different are cases like the noun *slím* which most commonly means ‘slime’ but is also found in the meaning ‘hindrance’, which seems to be of different origin and has given rise to a new homographic headword.

Words from related words (70 examples)

Another group of words has been labeled as originating from related words. This group is similar to the homograph group and most of the examples are explained in a similar manner as resulting from editing an already existing word. The related words are usually a different kind of word formation, either with a different suffix or different inflectional pattern. We can take as an example the neuter noun *tagl* ‘tail of a horse’ which has given rise to a new word *tögl*, a feminine noun that shows a different vowel development, but means the exact same thing. The only example of this new word was erroneously filed under the neuter noun, but after all the examples were categorized as part of the editing process the example was found to represent a different noun class that could only be accounted for with a new headword. This word was subsequently added to the lemma list.

Words added from similar words (21 examples)

Yet another group of words has been labeled as arising from graphically similar words, that are probably not related. The inception of these kind of words is very similar to the category above, except that the words are not related, e.g., the word *fírnska* is established in relation to the editing of a similar word *fíflska* ‘foolishness’ and seems to mean the same thing.

Words added from dissimilar words (5 examples)

In this small group we find a few words that have been added to the lemma list in relation to work on completely different words, where the dictionary editor has come across them. The only way to determine this for sure is when the editor in question has written a note about how this came about. For example, we find the word *vábein* of unclear meaning, where a note has been added that this word is established in relation to work on the word *kvikvendi* ‘living thing’.

Misplaced words (9 examples)

This is a small group of nine words that were added to the database after a batch of old citation slips was found in a drawer that had been used in conjunction with the editorial work on the first print volume. Most of these words are compounds where the word *altari* ‘alter’ is a member, e.g., *formessualtari*, *guðsmóðuraltari*. These slips were discovered by coincidence and seem to be a result of a filing error.

Newly excerpted words (46 examples)

This group mostly contains examples of words that have been discovered in new scholarly editions of lesser-known manuscripts. Even though most Old Norse prose texts have been published, there still remains a large body of manuscripts of various textual significance that has not yet been thoroughly accounted for. The words in this group are found mostly in small fragment texts that have recently been published, such as *hógværisandi* ‘spirit of modesty’ taken from a 2018 edition of a prayer text. In some cases, a new edition gives rise to a reclassification of a known example,

which requires the establishment of new headword, e.g., *huggøði* > *hugøði* in a new edition of a bishop's saga from 2018.

Unclear (42 examples)

Sometimes it is impossible to figure out why a particular word was added to the list of lemmas. It is likely that in most of such cases the dictionary editors have simply come across an interesting word in relation to their work on the rest of the vocabulary and subsequently have decided to add it to the database. It is therefore probable that most of these have a similar history as the group of words added in connection with work on unrelated words and should perhaps be counted with them.

6 Results

The results show that an overwhelming majority of the additions to the lemma list is a consequence of editorial work on other related or unrelated words where (re)evaluation of textual evidence has brought to light new independent headwords. In most cases, the additional lemmas are compounds that the editors came across when editing the simplex form of one of the members of the compound. Another significant contributive factor is the reconsidering of morphological forms and homographic variation. Only a small portion of the words in question are completely new words, which have been overlooked in previous lexicographic descriptions of the language or have been found in newly reevaluated text material.

7 References

- Bakken, K. (1995). *Leksikalisering av sammensetninger. En studie av leksikaliseringprosessen belyst ved et gammelnorsk diplommateriale fra 1300-tallet*. Avhandling for graden doctor artium. Oslo: Universitetet i Oslo.
- Bernhardsson, H., Haugen, O.E. & Berg, I. (2019) "Normalisation." Ch. 10 of *The Menota handbook: Guidelines for the electronic encoding of Medieval Nordic primary sources*. Gen. ed. Odd Einar Haugen. Version 3.0. Bergen: Medieval Nordic Text Archive. Accessed at: <http://www.menota.org/handbook.xml> [15/06/2021].
- Fritzner, J. (1886, 1891, 1896): *Ordbog over Det gamle norske Sprog 1–3, rev. udg.* Kristiania: Den norske forlagsforening.
- handrit.is*. An online catalogue over Icelandic and Nordic manuscripts. Accessed at: <https://handrit.is> [15/06/2021].
- Johannsson, E. (2019). Integrating analog citations into an online dictionary. In C. Navarretta, M. Agirrezabal & B. Maegaard (eds.) *Proceedings of the Digital Humanities in the Nordic Countries 4th Conference*. Copenhagen: University of Copenhagen, Faculty of Humanities, pp. 250-258.
- Johannsson, E. & Battista, S. (2014). A Dictionary of Old Norse Prose and its Users – Paper vs. Web-based Edition. In Abel, A. et al. (eds.) *Proceedings of the XVI EURALEX International Congress: The User in Focus, 15-19 July 2014*. Bolzano/Bozen: Institute for Specialised Communication and Multilingualism, pp. 169-179.
- Johannsson, E. & Battista, S. (2016). Editing and Presenting Complex Source Material in an Online Dictionary: The Case of ONP. In T. Margalitadze & G. Meladze (eds.) *Proceedings of the XVII EURALEX International Congress: Lexicography and Linguistic Diversity, 6-10 September 2016*, Tbilisi: Ivane Javakhishvili Tbilisi State University, pp. 117-128.
- Johannsson, E. & Battista, S. (2020). Brugere og forskellig anvendelse af ONP. In C. Sandström et al. (eds.) *Nordiske Studier i Leksikografi 15: Rapport fra 15. Konference om Leksikografi i Norden, Helsinki 4.-7.*, pp. 152-161.
- Jónsson, F. (1931). *Lexicon poeticum antiquæ linguæ Septentrionalis / Ordbog over det norsk-islandske skjaldesprog*. Original edition by Sveinbjörn Egilsson. 2nd ed. Copenhagen: Kongelige nordiske oldskriftselskab.
- ONP = Degnbol, H., Jacobsen, B.C., Knirk, J.E., Rode, E., Sanders, C. & Helgadóttir, Þ. (eds.). *Ordbog over det norrøne prosasprog / A Dictionary of Old Norse Prose*. ONP Registre (1989). ONP 1: a-bam (1994). ONP 2: ban-da (2000). ONP 3: de-em (2004). ONP Nøgle/Keys (2004). Copenhagen: Den Arnamagnæanske Kommission.
- ONP Online*. Ordbog over det norrøne prosasprog / A Dictionary of Old Norse Prose. Accessed at: <http://onp.ku.dk> [20/03/2021].
- Wills, T. & Johannsson, E. (2019). Reengineering an Online Historical Dictionary for Readers of Specific Texts. In Kosem, I. et al. (eds.) *Electronic lexicography in the 21st century: Smart lexicography: Proceedings of the eLex 2019 conference*. Brno: Lexical Computing CZ, pp.116-129.