
Pär Nilsson

REPORT ON THE REVISION OF THE SWEDISH ACADEMY DICTIONARY – AND THE SEARCH FOR “OLD NEOLOGISMS”

Abstract In 2023, the first edition of the diachronic-contemporary *Swedish Academy Dictionary* (SAOB) was finished after more than 130 years of work. The dictionary contains over 500,000 entries and describes the Swedish language from the 1520s to the present day. However, a consequence of the extended time of publication is that a significant part of the content, in particular in the first volumes, was out of date even by the time of completion of the full edition. Headwords had become obsolete, the meta-language was outdated in several regards, and a lot of modern words are lacking in the lemma list of the dictionary. To address this, a revision project was initiated in the beginning of 2024. This article reports on the revision in more detail, with emphasis on the supplement of new entries.

Keywords Swedish Academy Dictionary; historical dictionaries; dictionary revision; neologisms

1. Introduction

Up to 2023, the *Swedish Academy Dictionary* (SAOB) was sometimes said to be the longest ongoing non-finished dictionary project in the world. Correct or not, the project was initiated as early as in the 18th century, but it did not get its lexicographical breakthrough until the last decades of the 1800s when a permanent editorial board was installed in Lund. Under this new arrangement, the first volume of the dictionary was published in 1898 after 10 years of preparations (cf. Malmgren, 2011). Since then a new volume has been published on average approximately every three or four years. Of course, there were some worries about the pace of publication throughout the years, and more than once it has been questioned if the dictionary were ever to be finished (see e.g., Lundbladh (2003, p. 100) who refers to discussions on this topic in the first decades of the 1900s). However, since 2023 the first edition of SAOB is officially completed. By the end of that year the 39th volume, containing words beginning with the diacritic letters Ä and Ö, was printed and launched on the dictionary website <https://www.saob.se/>. But since the work has progressed only alphabetically forward, with no revision at all for older parts of the dictionary, a lot of entries, in particular in the first volumes, are out of date already, either because the words themselves have become obsolete or else the actual phenomenon that the notions denote. As the result of the long publication time many words and meanings, coined after the publication of the current alphabetic span, are also missing. In addition, some of the meta-linguistic terms and expressions used in the dictionary have become outmoded. Hence, there is room and reason for revision to be done. Therefore, a revision project was started at the beginning of 2024 with the goal of achieving a more balanced dictionary, without obvious gaps in the list of lemmas and with a modernized meta-language.

The aim of this paper is to describe the revision project and its content further, with the main focus on the addition of new entries and new headwords, i.e., the part of the dictionary work that has progressed the farthest at the time of writing. In order to grasp the nuances of this description more easily, a substantial background description of SAOB is also given. The paper is organized as follows: Section 2 describes the SAOB dictionary project, from a historical perspective as well as in terms of content, macro- and microstructural organisation etc. In Section 3, the general principles and different parts of the revision project are outlined. Section 4 deals with one of these parts in more detail, namely the neologisms (in SAOB terms), and in Section 5, lastly, some concluding remarks are made.¹

2. The Dictionary – a General Background

When the Swedish king Gustav III founded the Swedish Academy in 1786, it was declared in the Statutes (§ 23) that “it is [...] incumbent on the Academy to compile a Swedish dictionary”.² The following year the Academy members themselves initiated the dictionary work simply by distributing the letters of the alphabet within the group, after the model of the French Academy (Svensén, 2015, p. 87). For the Swedish Academy, however, this approach never led to a finished dictionary. As already indicated, the method of working did not find its form until an external editorial board with recognized lexicographical competence was installed in the 1880s. As a scientific leader of the project, Sweden’s most prominent lexicographer at the time, K. F. Söderwall, by then just in the process of finishing his solely historical dictionary of the Swedish medieval language (OSM), was chosen. Given the scope of his completed dictionary, it was decided that the period of description for SAOB was to begin where OSM ended, namely by the 1520s. (SAOB’s oldest source is the first volume of Gustav I’s registry from the year 1521, which thus marks the boundary between old and modern Swedish in the dictionary.)

The first task was to draw up a detailed plan of how the work was to be carried out and what source material would be selected and excerpted. This was followed by several years of collection before sufficient material was available in the archives to begin the first part of the first volume of the dictionary, which was eventually published in 1893 (the first full volume was completed 1898). A great deal of the original plan in terms of layout and content was then valid until the completion of the entire dictionary work. Svensén (2015, p. 100) points out, for example, that even current principles for the selection of compounds and derivatives or for technical terms were already specified in this plan.

2.1 Type, Target Groups and Organization of Entries

SAOB has a historical orientation. But unlike, for example, the previously mentioned Söderwall’s OSM, SAOB also has a contemporary perspective (according to Svensén’s (2009) typology, SAOB can be categorized as a diachronic contemporary linguistic

¹ Thanks to two anonymous reviewers for providing detailed and most helpful comments and suggestions.

² <https://sa-admin.lb.se/assets/3c46295e-685c-4bf1-a39d-61f1d8cd6f85.pdf>

dictionary; cf. Nilsson (2023, p. 25f) for a more complete typology of SAOB based on Svensen's terms). However, the contemporary perspective is relative – the contemporaneity in relation to later readers has, for obvious reasons, changed over the years – but the aim has always been to describe the words (their forms and sense development) from the 1520s to the current date.

The historical principle entails that the chronologically oldest meaning (globally speaking, not necessarily the first one encountered in Swedish) is presented first, that every lexical sense and micro-sense is supported by an authentic quotation, and that each meaning section in the dictionary is sampled by the oldest quote representing the current section. The same principle applies to subordinate meaning sections, whose first quote in order is the oldest found in the editors' collections.

Not only the perspective of time (diachronic-contemporary), but also the target group of users rests on two legs. The entries contain a lot of information and fairly advanced descriptions, but are still not aimed only at advanced users. In the preliminary preface, which was appended to the first half volume issued of the dictionary, it is stated: “that the dictionary is intended not merely for scientists, but for the entire knowledge-seeking Swedish public” (available: <https://www.saob.se/om>; my translation). This dual focus can still be said to apply, since the following is stated in the current description of the dictionary on the website [saob.se](https://www.saob.se): “Anyone interested in words and expressions will benefit from the dictionary, which is also an important aid in scientific research and a rich source for researchers or anyone who may want help when reading, in particular, older texts” (again in my translation).

The organization of entries, their macro- and microstructure, is clearly marked by its time of origin, both in terms of the printed medium (with a lot of space-saving principles) and the historical perspective, which applied to most, or all, linguistic work at the time the SAOB project in its current form was initiated (cf. Teleman, 2015). A complete account of lexical semantics and morphology, thus, more or less by necessity entailed a chronological account.

One important impact of these organization principles in SAOB, is that secondary lemmas, in terms of word-formation logic, are presented in subordinate divisions (subentries) under their respective primary headwords. Hence, the entry of simplex VAX (‘wax’) contains subentries like the compound VAX-DOCKA (‘wax doll’) and the derivative VAXIG (‘waxy’). The arrangement has space-saving reasons (the full account of forms and the comprehensive etymological derivation are given in the headword entry only and relevant parts are not repeated in the subentries), but is also grounded on the logic-chronological idea that a simplex word is more original than its compounds and derivatives (see Eaker (2006) for a closer description of SAOB's order of presentation). This kind of macrostructure is sometimes labelled *nest-alphabetical* (cf. Svensén, 2009). The term refers to the “nests” that are formed when morphologically related words (headwords, compounds, derivations, particle verbs, etc.) are brought together. The subordinated lemmas (the nest-lemmas) can in turn include their own compounds and derivatives, which gives the macrostructure a hierarchical character.

Also at the microstructural level, hierarchy is applied in the SAOB entries. A main sense (marked with an Arabic numeral) can have a subordinated sense (small italics) which in turn can have a subordinated sense (small Greek italics) that can have a subordinated sense (small Greek italics + prime symbol (')). The very first part of each entry, the head of the entry, consists of form and etymological descriptions. This section is often very dense in information (and full of space-saving abbreviations), particularly in headword entries. As mentioned earlier, the completed dictionary is published in print, but the corresponding content of the SAOB is also available in digital form. In Figure 1, the microstructure of a headword entry in the digital version is illustrated:

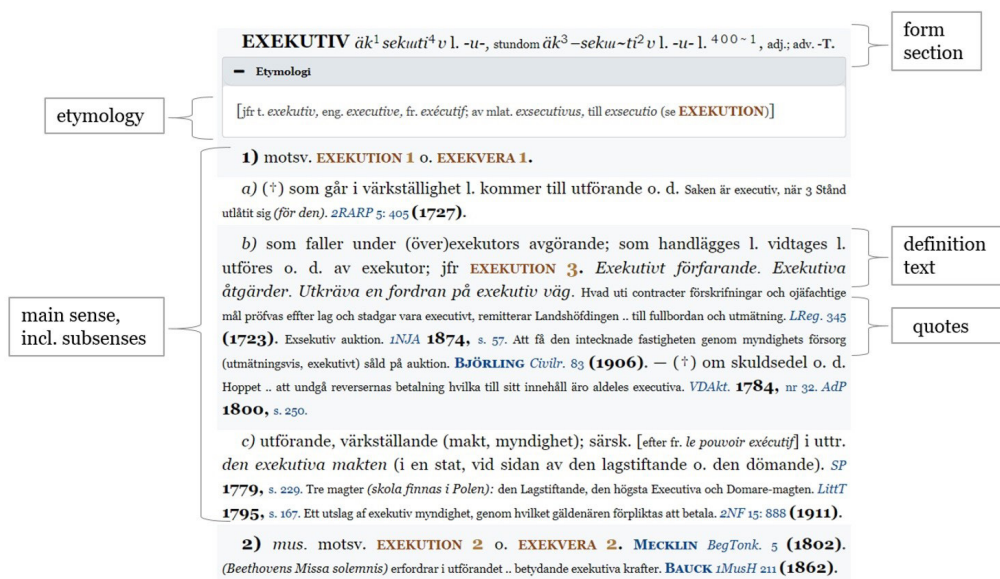


Fig. 1: Demonstration of SAOB's microstructure

2.2 Completion of the First Edition

As mentioned above, the work on the dictionary proceeded with a relatively similar result in terms of the entries' layout since the end of the 19th century and onwards. But in terms of detail and precision, and thus the speed with which the endpoint was approached, the work has changed character over the years. According to the original plan, the editorial board expected that their work would be completed – and that the first subscribers who signed up would see a complete work on their bookshelves – within a period of 10 – 12 years (Svensén, 2015, p. 99). However, the editors did not realize what an extensive task they had undertaken. In fact, this estimated time would have to be multiplied by 10; in total, the project lasted more than 130 years.

Several times the SAOB project has been threatened with closure, and the dictionary work has subsequently had to be reorganized more than once. Larsson (2014) identifies four periods in the history of SAOB that represent different levels of scope and ambition. In the fourth and final period (2000 – 2023), an approach was applied (as a reaction to the period before) that resulted in relatively concise and less detailed descriptions and definitions.

The completed dictionary comprises 39 volumes (including a half-volume with a list of the roughly 24,000 different sources cited in the dictionary) which include more than 500,000 lemmas distributed over more than 33,000 pages. As mentioned above, the dictionary is also available in an identical digital edition, which was first introduced in the 1990s but was re-digitized in the 2000s and is available, since 2016, on the website *saob.se*. Since 2017, the dictionary is also available on the dictionary portal *svenska.se* with its sister dictionaries SO ('The Contemporary Dictionary of the Swedish Academy') and SAOL ('The Swedish Academy Glossary'). The completion of the first volume of SAOB was officially announced in December 2023 at the Swedish Academy's Grand Ceremony in the presence of HRH King Carl XVI Gustaf.

3. Revision

Several years before the completion, it was clear that the dictionary would be in need of revision. In the last years of work, a revision project was therefore planned in parallel with the completion. Part of this preparatory work consisted of what was internally called "a revision tour". During this tour, universities and other research institutes were visited throughout Sweden (with stops in Lund, Växjö, Gothenburg, Stockholm, Uppsala, and Umeå) and in Finland (Helsinki). During the regular linguistic research seminars, various thematic aspects (one per institution) of an intended revision were discussed, and in the invitation to the seminar participants, the historically unique opportunity to influence the formation of SAOB's revised edition was highlighted. In addition, various aspects of the revision were presented and discussed at conferences and symposia in the Nordic countries, and a special half-day seminar with this theme, and with invited speakers, was arranged in connection with the conference *Svenska språkets historia* ('The history of the Swedish language') at Lund University in autumn 2021. During these seminars and presentations, topics such as grammar in the dictionary, material and sources, way of presentation, improvements of the microstructure, technical language and Finland-Swedish words and expressions (approx. 5 % of the population in Finland has Swedish as their first language) were discussed.

Many important points of view were collected, not least in terms of readability, gender and pronunciation designations. All in all, the editors got several clear and concrete ideas and more general input to serve as a point of departure for further work. However, to what extent these acquired insights and points of view could be applied in practice, was, of course, dependent on the financial framework.

3.1 A Brief International Outlook

There are close lexicographic equivalents to SAOB in some other European languages. In this context, *Deutsches Wörterbuch von Jacob und Wilhelm Grimm* (DWB), *Woordenboek der Nederlandsche Taal* (WNT), the *Oxford English dictionary* (OED) and *Ordbog over det Danske Sprog* (ODS) can be mentioned in particular. The first three of these were initiated before the SAOB project started. Hast (1983) claims that WNT and OED in particular served as role models for the SAOB editors at an early stage.

At the time of the completion of SAOB, the first editions of these sister dictionaries had long since been published. In several cases, supplements have also been added or other revision work has begun; see e.g., Scheider (2001) for a description of the revision of DWB, Kinable (2006) for revision aspects related to the handling of quotations in WNT, and Trap-Jensen (2019) of the retrodigitized ODS. SAOB's situation is thus, of course, not unique, even if the number of similar projects is limited. As in the initial stage of the project, the editors now have the opportunity to turn their eyes to these dictionary projects to learn lessons and examine how similar problems have been dealt with. Not least relevant for SAOB is, again, the work carried out by the editors of OED, both in view of the very similar micro- and macrostructures of the two dictionaries, and because editing systems from the same software company are used (namely IDM, cf. Section 3.3 below and see Nilsson (2024b) for a somewhat more detailed description). With regard to OED, cf. also Simpson (2020) and Elliott & Williams (2006).

3.2 Financing

Returning to the Statutes mentioned in Section 2, the Swedish Academy could be said to have fulfilled its task with the completion of the first edition of SAOB, and therefore they had no formal obligation to continue the publication. In addition, the Academy also publishes two other dictionaries (SO and SAOL, mentioned above, the latter of which was originally based on the same source material as SAOB). From the early 2000s, there was an economic agreement between the Swedish Academy and the Swedish state, which to a great extent financed the publication of SAOB. But this agreement ended with the completion of the first edition of the dictionary, and the Swedish Academy announced in 2021 that it could no longer afford publication. This was followed by a period of uncertainty regarding the future of SAOB, before the course turned again in 2022, when a collaboration was initiated between several different foundations which would enable the financing of a revision project corresponding to approx. 7 years of work starting in 2024. The Swedish Academy is still one of the financiers, but now together with the Helge Ax:son Johnson Foundation, the Society of Swedish Literature in Finland, The Royal Swedish Academy of Letters, History and Antiquities, and the Bank of Sweden Tercentenary Foundation.

3.3 A Three-Legged Revision Approach

Based on the thorough preparatory work, the editors had a relatively clear image of what is in need of update in the dictionary. But given the financial framework, the relevant questions to pose are: What can be done in this relatively limited period of time? How much is it possible to revise? The task is thus to accomplish as much as possible with as high standard as possible during the given period. This means, obviously, that not everything which is desirable to change can be changed during the project period. On the other hand, further improvements can possibly be made at a later stage, if the project is extended beyond these 7 years.

Early in the process, three distinct areas, or three different ways of approaching the revision, were identified. The main principles behind these approaches are summarized below.

Firstly, there are certain elements, or aspects, recurring in all or many entries that can be managed in more or less the same way at the same time. This group of changes, and the assignments related to it, is labelled *aspect revision*, and it includes heterogeneous types of aspects to deal with. A prerequisite for this task is the implementation of a new software (Digital Publishing System (DPS), provided by the French company IDM; cf. e.g., Arts, 2014; Kocjančič et al., 2006). The content of the previous digital edition has been transferred to this system and re-tagged. The result is a completely digital database, instead of the previously only digitized version of the printed text.

Optimally, the various aspects can be handled automatically by replacing one kind of element with another. But in practice, this is rarely the case, as most of the tasks require some form of manual effort. Most straight forward are tasks such as untangling abbreviations that are difficult to interpret, and replacing obsolescent spellings (which do not coincide with correct forms of other words) among lemmas and meta-language.

More demanding is the work of replacing the old gender designations from the first edition with modern equivalents (both the gender system itself and the meta-terminology have changed during the time of publication of SAOB). The issue here is that the detailed descriptions (often including several different fonts) are difficult for the software to decode, and more advanced descriptions have to be translated manually. However, many gender designations are more concise and can be translated in a satisfactory way automatically. The editors' approach here is to prioritize the unproblematic instances and make the entries in which they are found as up-to-date as possible in this respect, and then deal with the more complicated cases later, even if this implies that SAOB will, for a period of time, contain entries with two different systems of denoting gender (one based on the older system's combinations of grammatical and sexual gender, and one based on grammatical gender alone; see Nilsson (2024b) for an account of SAOB's way of dealing with gender in the first edition and in the revised version).

More or less on the same level of difficulty is the task of rearranging the content of the definition text based on the hierarchy that already exists in the entries; i.e., SAOB has a subtle way of separating different sections within a definition using punctuation marks and certain keywords. A semicolon plus the word "särskilt" ('especially') indicates that what follows constitutes a subordinated special use of what was previously said, while a semicolon plus "älv." ('also') refers to a more independent secondary sense. By automatically dividing the definition text with a new line for each semicolon and also a short indent for descriptions beginning with "särskilt", the hierarchy can be highlighted and clarified, and at the same time the text becomes less compact. This can be (and has in fact been) done automatically. But a complicating circumstance for the machine interpretation is that the key words "särskilt"/ "älv." can be modified and combined in various ways (e.g., *formerly also*, *sometimes also*, *except*

for older conditions now especially of ..., etc.) and semicolons can (with or without difference in meaning) be replaced by commas or parentheses. This aspect, too, can thus be classified as semi-automatic.

Even more demands on manual effort are required by the task that has to do with the description of semantic relations (see Nilsson, 2024a). Labels such as *transferred use* have been consistently used in the dictionary, but partly in different ways, which will now be adjusted. Such labels can be easily identified in the editing system, but replacing an inappropriate label requires a certain amount of manual work and individual assessment (based on certain criteria).

Secondly, and as a means of compensating for the deficiencies of the aspect revision, *manual revision of the individual entries* is needed. This part of the revision project is still in development. At the time of writing, a pilot study is underway. Manual efforts can of course be made in different ways, and possibly several different methods will be applied later. But since many entries contain a lot of old-fashioned or obsolescent elements – both linguistically (especially in the meta-language; cf. e.g., how plural verb conjugation occurs in SAOB entries despite the fact that the usage disappeared in Swedish in the middle of the 20th century) and regarding the phenomena described (in the entry BÄVER ('beaver') this species is described as extinct in Sweden, which is no longer the case) – it is urgent not to concentrate on a few selected single entries, but to review as much of the dictionary's content as possible. Hence the strategy at present is to focus on the low-hanging fruit, on as wide a range as possible, both in terms of the number of entries and the different types of updates within each entry.

Naturally, the older entries are in most need of revision, and therefore entries with initial letters from A–R (written between the 1890s and 1960s) are selected in this respect. The aim is to go through all entries but with a light hand. Each entry is revised in two levels and by two different editors – one editor reads the whole text and changes outdated wording and inaccuracies in the entries and adds new senses (if there is an obvious need). More complicated cases are skipped for the time being, which means that restructuring of large entries is not done at this stage. The task of the other editor is to supplement entries with modern quotations. This is a relatively uncomplicated task, but one that affects the reader's impression quite a lot. Often it is the lack of examples in a modern context that makes words described in the 19th century perceived as outdated, and just the addition of citations from the 21st century can thus significantly modernize an entry. At the moment, the editors work their way through the alphabet in pairs, with the equivalent of 100 columns in the printed dictionary at a time, which seems a manageable amount. The two editors then review each other's work. It should be added that in the manual revision the chance to adjust the outcome of the aspect revision is also given, which makes this task a kind of control station.

Thirdly, SAOB's long publication time has implied that many modern words are missing, especially in the first half of the dictionary. An area in obvious need of revision is, thus, the *addition of new entries*, the part of the project that is internally called *nyord*, 'neologisms', but whose definition in this case differs slightly from many

other contexts (see further next section). The work related to new entries is a separate track in relation to the other two revision approaches, in that it does not involve the existing entries to any significant extent. The task is more straightforward in the way that the old, retagged structure does not have to be taken into account (except in the case of cross-references). Instead, the content of the entries can be written directly into the purpose-designed software tool.³

Overall, the addition of new entries is the one of the three tasks in which the most time will be invested. In the current project plan, it is estimated that about 50% of the resources (it should be said that the editorial staff currently consists of 17 people, with different employment rate) will be spent on this work, and the remaining 50% on the other two tasks described above (including IT developments).

The assignment of finding, collecting and defining these neologisms are described in more detail in the next section.

4. New Entries – and the Search for “Old Neologisms”

As has already been stated, new entries are needed in SAOB, and this need is not evenly distributed throughout the dictionary. The lack is even more obvious in the first volumes, or rather in the first half of the dictionary, partly because the excerption of source material to a large extent was done in advance, but also because several entries were published at the end of the 19th and at the beginning of the 20th century. Therefore, also in this part of the work, the focus is on words with initial letters from A–R. Based on these SAOB’s needs, a neologism in this context is not to be understood as a lexical unit perceived as new or recently established or not yet (fully) lexicalized (see definitions discussed by Klosa & Lungen, 2018, pp. 560–661; cf. also Klosa-Kückelhaus & Kernerman, 2023). The perspective here is slightly different since the concept of *new* is relative, and in the context of SAOB’s long time span, it aims at something else than, for instance, in the latest edition of a modern contemporary dictionary. To address this relativity, neologisms, or the new features, are here simply related to the lemma list of SAOB. It refers to a headword that the dictionary will be supplemented with, because it is considered a deficiency (due to SAOB’s aim to reflect the language in the most representative possible way) that it is missing in the first edition. A definition thus more in line with the one mentioned by Afentoulidou & Christofidou (2021, p. 599): “any lexical unit being classified as a member of the active vocabulary of the speech community [...], based on certain criteria, which would allow its inclusion in general-purpose dictionaries.” (More specific criteria are commented below.)

A relevant question to ask before the addition of new entries in SAOB’s revision project is how many headwords should be added. How much can be done, and what is a sufficient amount to make the dictionary better and to overcome the biggest flaws? There is, of course, no clear answer to these questions, but the editors have come to

³ The development and design of the application of the software will not be discussed in more detail in this paper. But it must be emphasized that the adaptation to the structure of SAOB is a huge effort that has been going on for several years.

the conclusion that 10,000 lemmas, corresponding to about one printed volume, is a reasonable number, and that is what is aimed for. The estimate is based both on how much of the total amount the supplement makes up, but of course also on how much time of work will be spent.

Although the first edition of SAOB contains a respectable number of entries, more than any other Swedish dictionary, it is not a difficult task to find relevant words to add to the dictionary. On the contrary, the challenge is to restrict the selection and to compile narrow criteria for inclusion.

4.1 Sources for Neologism Candidates and Specified Selection Criteria

First of all, it must be decided what sources to use to find neologism candidates. In the first edition of SAOB the framework was fairly fixed, and, more or less, the later editors came to a set table; an excerpted, physical source material⁴ was available and the words contained therein ended up in the dictionary mainly if they met the following criteria: they were not proper names, they were not exclusive technical terms or dialect words, and they were not purely occasional formations. The vast majority of simplex words in the material were thus included, in addition to a rich selection of compounds, derivatives and particle verbs. Now the situation is different, and input can be obtained from almost anywhere. The aim will still be to describe the Swedish language in as comprehensive a way as possible, and the criteria described above still apply. Five different channels for accessing neologism candidates have been narrowed and will be used, of which the work so far is based on one of them.

The first source of candidates is found in-house. Since the excerption of the printed sources used for the first edition was executed continuously from the preface (mentioned above) to the 1990s on words in all the alphabet, though some of the letters were already treated in the dictionary, a certain part of the material labeled *late incoming* is available in the archive of the dictionary, which has not been utilized so far. This partial material consists of approx. 2–300,000 excerpts. For the sake of candidate selection, the lemmas in this material are being manually excerpted and listed. The compiled lemma list from this source material is later to be matched and aligned with the lemma list of SAOB. Samples show that relevant candidates can be found this way.

As mentioned in Section 2.2, SAOB is digitally available together with the two other dictionaries of the Swedish Academy, SO and SAOL, on the site <https://svenska.se/>. Also their lemma lists are matched with SAOB's. The result shows that about 40,000 lemmas found in SO and SAOL are lacking in SAOB within the relevant alphabetic span. The SAOB editors have gone through this list in pairs and have selected just over 3,000 relevant words to include in the revised version of the dictionary. This selection forms the basis of the work on new entries that has begun, and which at

⁴ SAOB's source material consists of approx. 8,2 million excerpts collected from about 20,000 sources of different genres. See e.g., Svensson (1992) for a discussion on SAOB's material with emphasis on the representativeness of the sources; cf. Malmgren (2011) for a shorter description in English.

the time of writing has arrived at the last words beginning with the letter B in the alphabetical order. (How the selection was carried out is described in more detail below.)

In addition, a targeted corpus has been compiled for the purpose of the dictionary. This corpus, called SAOB1950, is freely available on Språkbanken text's search tool *Korp* (for more information about this tool see e.g., Borin et al., 2012). The corpus contains material from fiction and non-fiction books collected from the university library in Lund with the intention of balancing the newspaper text bias that is present in other available (Swedish) corpora. The books are published, as the name of the corpus suggests, from 1950 onwards and they are selected to cover several different genres, both by direct selection (certain genres and subgenres) and by random selection (every *n*th book on every *n*th shelf). The corpus is relatively small, about 50,000,000 tokens, but of high quality for the dictionary purposes. The total amount of tokens in the corpus will be matched with SAOB's lemma list. Samples show that this source will also be useful. It contains a slightly different kind of words than the aforementioned sources.

Except for these sources, candidates are collected manually by listing lemmas that are relevant to add in connection with other neologisms being added. For example, the definition of the neologism compound *afterski band*: "that typically plays covers", resulted in the word *cover* ending up on the list of neologism candidates. In addition, candidates can be collected through the log files of so-called *not-found*s, i.e., frequently recurring non-successful searches of words that are not included in the dictionary (cf. Bäckerud et al., 2021). The point of having several different channels to collect candidates is, of course, that a more balanced influx and representation from different areas can be accomplished. However, in this case, more emphasis is placed on quality than quantity. What is decisive is the composition and different nature of the sources.

The following criteria (in decreasing order of priority) form the basis for the assessment of candidates.⁵ Words that are given priority:

- belong to (or in an earlier language stage have belonged to) almost all language users' (at least passive) vocabulary; for example (here and henceforth, if no translation is given: the Swedish equivalents of): *abstinence*, *acetone*, *adhd*, *airbag*, *activate*, *allergy*, *avocado*;
- can be considered somewhat specialized, often belonging to a more or less technical language, but which quite a few language users may need to acquire through practical needs or leisure interests or community involvement or encountering illness and the like; examples: e.g.: *agility*, *antonym*, *arbitrage*, *assist* ('scoring pass'), *audiologist*;
- may be considered "difficult" or "learned" words that should be explained in a dictionary (even if they have low frequency), e.g., *annihilate*, *adiaphora*, *arborist*;

⁵ The candidate collection based on SO's and SAOL's lemma lists was performed in pairs letter by letter. The words that both editors in the pair had selected ended up on the final list of neologism candidates.

- occur in the meta-text of SAOB (or are found in the list of lemmas as the second part of a compound); e.g.: *anion* (in the entry CATION), *anträffbar* ('accessible, available', in the entries O-ANTRÄFFBAR, SVÅR-ANTRÄFFBAR).

Simplex words are prioritized, not least since it is unproblematic for the editors during the actual editing process to supplement the simplex entry with one or more nest-lemmas if needed. But also (frequent) compounds and derivatives are collected if they are:

- lexicalized in an opaque special meaning (at least in one of its meanings), because the relationship between the parts is completely unique, or one or both parts as a simplex word exhibits such extensive polysemy that several different "meaning sums" are plausible for the compound; for example: *aha-experience*, *daghem* ('day care');
- more or less semantically transparent, but have nevertheless established themselves as "just what it is called" despite fully possible alternative expressions; that is, they are unpredictable from a production point of view; for example: *allergic shock*, *alliance free*, *atomic bomb*.

Other words are to be left unmarked as instances of a low-priority lack in the list of lemmas. However, even with the specified criteria, there is an imminent risk of the selection becoming too abundant. Therefore, some priority guidelines have been formulated. Firstly, names of varieties and sorts and name-like designations (which are not highly frequent or culturally central) are de-prioritized. This means for example that names of food dishes or all dog and cat breeds are not prioritized. Neither are all designations for inhabitants and languages collected. In fact, the aim is not to prioritize complete descriptions of different systems just for the sake of systematicity. A more pragmatic approach must be applied. Instead, emphasis is placed on frequency, relevance and cultural centrality. Secondly, the fact that SAOB has a historical claim is highlighted. In other words, slightly older words that are at risk of being phased out from the purely contemporary dictionaries are given high priority in SAOB. This means that words such as *bakelite*, denoting a material that is hardly used nowadays, *zairisk* ('Zairian'), despite being a designation for nationality, but in the capacity of a no longer officially valid term, the above-mentioned *daghem*, officially replaced by *förskola* ('preschool'), and the 1970s-scented *barkök* ('kitchen bar'), will end up high on the list. Thus, the task to large extent is to search for "old neologisms".

4.2 Editing and Reviewing

The work on adding new entries has at this point made some progress. So far, words like *acne*, *app* and *baseball* and such previously mentioned words as *airbag*, *activate*, *allergy*, *bakelite*, and *barkök* etc. have been added. (However, these entries have not yet been published online – a first round is expected to be published in 2025.) But it is not only the headwords themselves that are new, also the principles for describing the words and reviewing the written descriptions have several new elements. In this section, this news will be briefly discussed.

As mentioned in the previous section, there are a number of channels to identify neologism candidates. In the analysis and description of the individual word, however, the editors use more resources than these. As highlighted in Section 2.1, the first quote or sample of a form or meaning is important in SAOB. To access these, the editors use several different corpora plus the physical material available at the editorial office. Of these corpora, special mention can be made of the newspaper database of the National Library of Sweden, *Svenska Tidningar* (available at: <https://tidningar.kb.se/>), where a first quote is often found. In this respect, of course, SAOB's own corpus SAOB1950 is also important. In connection with the previously mentioned revision tour, many comments were received about the complexity of SAOB's definition text. As mentioned in Section 3.2, systematic work is being done to make the hierarchy in the definitions clearer. In addition, the neologism editors now have the opportunity to express themselves in a more accessible way in the definitions, which, for example, means that they can now avoid long (preceding) noun phrases (in favor of relative clauses) and complicated structural solutions, for instance, requiring parentheses within parentheses. In addition, typicality of uses and attributes, etc. can be made explicit with expressions such as: *typically of X or Y*, or *for instance in terms of Z* etc.

The editors are currently working with a span of 25 neologisms at a time. Information about all sources and citations used in each span of entries are saved in separate documents. When the first draft is completed, a colleague reviews the entries for all content (including cited sources). After further processing, the entries are reviewed by two other colleagues, before it is proofread and ready for publication. At the time of writing, the work has not yet reached this point, but with the preparatory work done, the route is thus marked out.

5. Concluding Remarks

In sum, SAOB is facing both exciting and challenging times. Everything can be questioned and much can be improved, and the possibilities are in many respects unlimited, for better or worse. It is therefore important to take a pragmatic approach, and the editors must in some respects apply a trial-and-error method in the initial revision work. After fairly thorough preliminary efforts, however, there is a stable basis to start from, and the fact that the dictionary will henceforth be continuously updated implies that mistakes, errors and inadequacies can be corrected step by step as they are discovered.

The various parts of the revision are made clear and manageable to an increased extent, not least when it comes to work with neologisms. Obvious gaps and deficiencies in the lemma list are identified, and the intention is not to aim at so-called comet words or the very latest novelties in the vocabulary, there are better dictionaries for that purpose. It thus seems appropriate for a diachronic-contemporary dictionary like SAOB to set the sights on old neologisms.

References

- Afentoulidou, V., & Christofidou, A. (2021). It's a long way to a dictionary: Towards a corpus-based dictionary of neologisms. In Z. Gavriilidou, L. Mitits, & S. Kiosses (Eds.), *Lexicography for inclusion: Proceedings of the XIX Congress of the European Association for Lexicography* (pp. 597–606). Democritus University of Thrace.
- Arts, T. (2014). The Making of a Large English-Arabic/Arabic-English Dictionary: the Oxford Arabic Dictionary. In A. Abel, C. Vettori, & N. Ralli (Eds.), *Proceedings of the XVI EURALEX International Congress: The user in focus* (pp. 109–123). Eurac.
- Borin, L., Forsberg, M., & Roxendal, J. (2012). Korp: The corpus infrastructure of Språkbanken. In N. Calzolari, K. Choukri, T. Declerck, M. Uğur Doğan, B. Maegaard, J. Mariani, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of LREC 2012* (pp. 474–478). ELRA.
- Bäckerud, E., Nilsson, P., & Sköldbberg, S. (2021). Så används Svenska Akademiens ordböcker på nätet. Implicit och explicit feedback från användarna. In C. Sandström, U.-M. Forsberg, C. af Hällström-Reijonen, M. Lehtonen, & K. Ruppel (Eds.), *Nordiska studier i lexikografi 15* (pp. 91–101). Nordiska föreningen för lexikografi.
- DWB: *Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm, digitalisierte Fassung im Wörterbuchnetz des Trier Center for Digital Humanities, Version 01/23*. Retrieved July 27, 2024, from <https://woerterbuchnetz.de/?sigle=DWB#0>
- Eaker, B. (2006). Historisk lexikografi i praktiken. Om momentordningen i SAOB. *LexicoNordica*, 13, 41–54.
- Elliott, L., & Williams, S. (2006). Pasadena: a new editing system for the Oxford English Dictionary. In E. Corino, C. Marelllo, & C. Onesti (Eds.), *Atti del XII Congresso internazionale di lessicografia/ Proceedings of the XIIth Euralex International Congress* (pp. 257–264). Edizioni dell'Orso.
- Hast, S. (1983). Om SAOB:s material. *Nyvenska studier*, 6, 159–192.
- Kinable, D. (2006). Computerized restoration of historical dictionaries: uniformization and date-assigning in dictionary quotations of the *Woordenboek der Nederlandsche Taal*. *Journal of literary and linguistic computing*, 21(3), 295–310.
- Klosa-Kückelhaus, A., & Kernerman, I. (2023). Unregistered words, neologisms, and dictionaries: An introduction. *Lexicography: Journal of Asialex*, 10(2), 87–93.
- Klosa, A., & Lungen, H. (2018). New German Words: Detection and Description. In J. Čibej, V. Gorjanc, I. Kosem, & S. Krek (Eds.), *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts* (pp. 559–569). Ljubljana University Press.
- Kocjančič, P., Krek, S., & Climent, P. (2006). Bilingual Dictionaries and IDM DPS: The Development of a Corpus-driven Slovenian-English Pocket Dictionary and its Implementation in the IDM Dictionary Publishing System. In E. Corino, C. Marelllo, & C. Onesti (Eds.), *Proceedings of the XII EURALEX International Congress* (pp. 431–438). Università di Torino.
- Larsson, L. (2014). En “mer l. mindre stor” stor ordbok – om variationerna i SAOB:s omfattning och ambitionsnivå. *LexicoNordica*, 21, 61–80.

- Lundbladh, C.-E. (2003). Kritiken av SAOB. *LexicoNordica*, 13, 273–284.
- Malmgren, S.-G. (2011). Das schwedische Akademiewörterbuch und das Wörterbuch der dänischen Sprache. In U. Haß (Ed.), *Grosse Lexika und Wörterbücher Europas. Europäische Enzyklopädien und Wörterbücher in historischen Porträts* (pp. 301–315). de Gruyter.
- Nilsson, P. (2023). *Lexikal betydelseutveckling i teori och praktik: en analys av fem definitionsformler i Svenska Akademiens ordbok och de semantiska förändringsmekanismer som de beskriver*. [Doctoral dissertation, Lund University].
- Nilsson, P. (2024a). Figurative, transferred or extended use? The use of semantic labels in the first edition and the revised version of the Swedish Academy Dictionary. In K. Š. Despot, A. Ostroški Anić, & I. Brač (Eds.), *Lexicography and Semantics. Proceedings of the 21st EURALEX International Congress* (pp. 123–131). Institute for the Croatian Language.
- Nilsson, P. (2024b). Glancing back, looking forward: Swedish Academy Dictionary completed after 130 years – now time for revision. [Manuscript in preparation]
- ODS: *Ordbog over det danske Sprog*, grundlagt af Verner Dahlerup, udg. af Det Danske Sprog- og Litteraturselskab, bd 1–28, København 1918–1956. Retrieved July 29, 2024, from <https://ordnet.dk/ods>
- OED: *The Oxford English Dictionary*, second edition, vol. I–XX (of the corrected reissue (1933) with an Introduction, Supplement, and Bibliography of *A New English Dictionary on Historical Principles*, vol. I–X, ed. by James A. H. Murray et al. (1884–1928) and Supplement, vol. I–IV, ed. by R. W. Burchfield (1972–1986)), Oxford 1989. Retrieved July, 29, 2024 from <https://www.oed.com>
- OSM: Söderwall, K.-F. (Ed.) (1884–1918). *Ordbok öfver det svenska medeltids-språket*. Retrieved July, 29, 2024, from <https://spraakbanken.gu.se/resurser/soederwall>
- SAOB: *Svenska Akademiens ordbok* (1898–2023). Glerups. Retrieved July 29, 2024 from <https://www.saob.se/> and <https://svenska.se/>
- SAOL: *Svenska Akademiens ordlista* (14th edition). (2015). Retrieved July, 29, 2024 from <https://svenska.se/>
- Scheider, M. (2001). Die Neubearbeitung des ‚Deutschen Wörterbuchs‘ von Jacob Grimm und Wilhelm Grimm. *Zeitschrift für Germanistik*, 11(3), 600–607.
- Simpson, J. (2020). *Preface to the Third Edition of the OED*. Retrieved July 29, 2024 from <https://www.oed.com/information/about-the-oed/history-of-the-oed/preface-to-the-third-edition-of-the-oed/?tl=true>
- SO: *Svensk ordbok utgiven av Svenska Akademien* (2nd edition). (2021). Retrieved July 29, 2024, from <https://svenska.se/>
- Svensén, B. (2009). *A handbook of lexicography: the theory and practice of dictionarymaking*. Cambridge University Press.
- Svensén, B. (2015). *Snille och smak. Svenska Akademien förr och nu*. Norstedts.

Svensson, L. (1992). Om SAOB:s material och en bibliografisk databas. In R. Vatveldt Fjeld (Red.), *Nordiske studier i leksikografi 1. Rapport fra konferense om leksikografi i Norden* (pp. 369–379). Nordiska föreningen för lexicografi.

Statutes of the Swedish Academy (1786). Retrieved July, 29, 2024, from <https://sa-admin.lb.se/assets/3c46295e-685c-4bf1-a39d-61f1d8cd6f85.pdf>

Teleman, U. (2015). *Nordister och nordistik i Lund – från 1860- till 1930-tal*. Lundastudier i nordisk språkvetsenskap A 73. Lund.

Trap-Jensen, L. (2019). State-of-the-art on monolingual lexicography for Denmark (Danish). *Slovenščina* 2.0, 7(1), 1–12. <http://dx.doi.org/10.4312/slo2.0.2019.2.1-16>

WNT: *Woordenboek der Nederlandsche Taal*, deel I–XXIX + Bronnenlijst (1943, 1966) + Supplement I A–Azuursteen (1956), 's-Gravenhage & Leiden 1882–1998. Retrieved July 29, 2024, from <https://gtb.ivdnt.org/search/?owner=wnt>

Contact information

Pär Nilsson

Svenska Akademiens ordboksredaktion

par.nilsson@svenskaakademien.se