

Louise Holmer, Ann Lillieström, Emma Sköldberg, and
Jonatan Uppström

TIME TO SAY GOODBYE REVISITED

On the Exclusion of Headwords from the Swedish Academy Glossary (SAOL)

Abstract In the revision process of dictionaries, adding new headwords or new senses to already existing headwords is what typically receives the most attention. In this article, we bring into focus the intriguing dilemma of exclusion of headwords from the Swedish Academy Glossary (SAOL), which is still published in print versions. In the e-dictionary-era, removing headwords may seem questionable. SAOL is, however, a contemporary dictionary which aims to reflect present-day Swedish. In order to keep the lemma list up to date, new headwords are added and obsolete words are removed. The editors of SAOL have practised lemma exclusion in connection with the revisions of new editions for almost 150 years. In this paper, we present SAOL and argue that lemma exclusion is crucial to SAOL's aim and target group. We also present our most recent corpus material, methods and tools included in this process.

Keywords lemma exclusion; SAOL; Swedish monolingual dictionaries; Språkbanken Text

1. Introduction

When revising dictionaries, the revision of the lemma list and the inclusion of new headwords in the dictionary naturally gain a lot of attention. The corpus material, the methods, and the motivation behind the selection of headwords (lemmas), are also a topic for various investigations.

However, the exclusion of headwords from dictionaries is not as often a topic of discussion in metalexigraphy. One reason for this may be the difficulty in detecting words that are falling out of use – especially in comparison to newly coined words. The declining use of a word seems to be a slower process compared to the formation and spread of a new word. In addition, a word that has previously shown a tendency to fall out of use may return with renewed force (which seems to be the case for the Swedish verb *prokrastinera* ‘procrastinate’). Hence, as Gellerstam (2005, p. 133) points out, the exclusion of lemmas is a delicate matter in practical lexicographical work (cf. Diamond, 2016). Gellerstam also states that, as a result, obsolete words often remain in dictionaries as “silent witnesses from another time,” possibly with a usage label such as “uncommon”.

Nowadays, the possible deletion of lemmas might also be considered irrelevant. For e-dictionaries, space restrictions do not apply in the same manner as they do, or previously did, for print dictionaries.

The aim of the present study is to analyse the list of headwords in the Swedish Academy Glossary (SAOL, 14th ed., 2015) and to investigate how (in)frequently inflectional forms of each headword occur in a large, contemporary text corpus. The motivation behind this is that low frequency in present-day texts is the main reason for exclusion of lemmas from SAOL.

The paper is structured as follows: first, we give a brief overview of the dictionary in focus, SAOL, with special attention to its list of lemmas and the principles for inclusion and exclusion of headwords. In section 3, we present the methodology and material in our study. In section 4, the results of the investigations are presented and discussed, and finally, in section 5, we relate the results to the lexicographical challenges concerning space restrictions in a dictionary that is published in print and in various e-versions.

2. The List of Headwords in SAOL

SAOL plays an important role in Swedish linguistic infrastructure (Holmer et al., 2024). It is a monolingual dictionary of present-day Swedish, focusing on orthography and inflection, but not semantics or phraseology. It is aimed at supporting production-related tasks and may also provide information on word formation such as linking morphemes in solid compounds (Berg & Malmgren, 2010). Moreover, SAOL is often regarded as an unofficial norm for spelling (Holmer et al., 2015; Sköldberg et al., 2019) and it is being widely used in schools, by teachers, in university courses, by linguists, by writers and by officials writing public documents.

There are several other Germanic languages with dictionaries similar to SAOL, for example, Duden's *Die deutsche Rechtschreibung* for German, *Woordenlijst Nederlandse Taal* for Dutch, *Retskrivningsordbogen* for Danish, *Tanums store rettskrivningsordbok* for Norwegian (Bokmål) and *Stafsetningarorðabókin* for Icelandic (see Gellerstam, 2009b; Vikør, 2009; Berg et al., 2010). They all have a tradition of regular revisions, and they were all published in print first, and have later, in various ways, been developed into different e-versions. See Figure 1 for an example of the print version of SAOL 14 (2015), with lemmas starting with the letter R.

re.kond.it.ion.era v. ~de ~t • åter försätta i god kondition	re.kord jämn adj. ~t ~a • till jämn 2	rekryt [rekry't] s. ~en ~er • soldat under grund- utbildning
re.kond.it.ion.er.ing s. ~en ~ar	re.kord kall adj. ~t ~a	rekryt.era v. ~de ~t • skaffa nya medlemmar el. pers- onal; nyanställa
re.kon.stru.era v. ~de ~t • återställa i ursprunglig form; ombilda; jfr rekonstruktion	re.kord lopp s. ~et; pl. ~	rekryt.er.are [-e'r-] s. ~n; pl. ~, best. pl. <i>rekryterarna</i>
re.kon.strukt.ion s. ~en ~er • till <i>rekonstruera</i>	re.kord låg adj. ~t ~a • r. ränta	rekryt.er.ing s. ~en ~ar
re.kon.strukt.ionsplan s. ~en ~er	re.kord lång adj. ~t ~a	rekryt.er.ings bas s. ~en ~er
	re.kord många pron.; pl.	rekryt.er.ings be.hov s. ~et; pl. ~
	re.kord mäss.ig adj. ~t ~a	
	re.kord nivå s. ~n ~er	
	re.kord not.er.ing s. ~en ~ar	
	re.kord prest.at.ion s. ~en ~er	

Fig. 1: Example from SAOL 14 (2015) in print, the letter R

As seen in Figure 1, SAOL lists lemmas, parts of speech and inflections in abbreviated forms for each headword, and in some cases short semantic descriptions. Solid compounds ending with a polysemous constituent, such as *jämn* in *rekordjämn* (lit. ‘record even’) (in the middle column), are followed by a reference to the relevant sense of the constituent (here *jämn 2*). The lemma *rekonstruktionsplan* (‘reconstruction plan’) (in the first column) illustrates the linking morpheme *-s-* between the constituents *rekonstruktion* and *plan*.

SAOL has about 126,000 lemmas distributed over the following major parts of speech:

- nouns: 94,000
- adjectives: 16,500
- verbs: 10,500
- adverbs: 1,200
- others: 3,800

Neologisms in Swedish tend to be nouns, often in the form of solid compounds. As a natural effect, candidate words for inclusion in SAOL are, therefore, often solid compound nouns; this also applies to candidates for exclusion (Berg et al., 2010).

SAOL has a long tradition in Sweden – its first edition was published in 1874 and it has since been fully revised and published in 14 editions over 150 years. From the first edition’s approximate 34,000 headwords to the latest edition and its 126,000 headwords, many new words have been included in the dictionary, but a fair share of obsolete ones have also been excluded (Holmer et al., 2024).

2.1 Lemma Inclusion in SAOL

As a rule, lemma inclusion criteria in SAOL involve frequency in written texts, a widespread use geographically, and preferably, use in texts that are aimed at the general public and non-specialists. Traditionally, the Swedish corpus material has mostly consisted of newspaper texts, but over time, the editors have also started to include other text types as well, e.g., social media (although edited texts are more relied upon than non-edited texts). See e.g., Gellerstam (2009a) for an in-depth study of the lemma lists in different editions of SAOL, and Berg & Malmgren (2010) for a discussion focusing on the inclusion of solid compounds in the glossary.

There are various aspects of lemma inclusion, and it is, in this study, impossible to give a satisfying overview of this field of research. Some (of many) recent studies include Klosa & Lungen (2018) on the methods of semi-automatic detection of new German words and their possible inclusion in the *Neologismenwörterbuch*. Another example is ten Hacken (2020) who discusses the relationship between new lemmas, the empirical reality, and the norms in general. Moreover, Sánchez & Maroto (2021)

provide a comparison between a more theoretical approach and speakers' insights in relation to new Spanish words.

When new editions of SAOL have been released (about every 10th year) the number of included words such as neologisms has received a considerable amount of attention in the media and among users. When it became known that the (then) forthcoming edition of SAOL 14 was going to include the gender-neutral pronoun *hen* in 2014, even international media wrote articles on the topic.

2.2 Lemma Exclusion in SAOL

A previous study of SAOL, named “Time to say goodbye”, focused on the exclusion of mainly solid compounds (Berg et al., 2010). It showed that about 10% of the headwords of SAOL 13 (2006, comprising about 125,000 headwords) were not in use in the reference corpus of about 250 million words (at that time considered a large corpus) of mainly newspaper texts – in spite of the fact that about 5,500 lemmas had been removed in relation to the previous edition.

An earlier study (Gellerstam, 2005) also showed similar results; about 10% of the lemmas of SAOL 12 (1998, with about 120,000 headwords) were not in use in a corpus of about 90 million words. It is important to note that irrespective of the number of headwords included in SAOL, the glossary will still only include a sample of contemporary Swedish vocabulary, since the possibilities of coining new compounds and derivatives are almost infinite (Berg et al., 2010; cf. Hanks 2019).

In an article by Svensson (2019) in *Språktidningen*, a Swedish popular science magazine, the author lists and comments on 1,001 headwords in SAOL 14 that may risk being excluded in the next revision due to low frequency in texts. Svensson asks the readers to adopt and use their favourite of these obsolete words in order to save them from endangerment. This campaign in *Språktidningen* was well received in the media, and also led to considerably more e-mails from the general public to the editors of SAOL.

The next edition of SAOL is planned for late 2025, and, as already mentioned, it will be published not only digitally but also in print. There is an explicit request from both members of the Swedish Academy and the SAOL editorial staff that the different versions of the glossary should include the same lemmas. In order to keep the print version accessible and in one volume, it is necessary to delete obsolete words and senses, although the digital version allows an expanded lemma list. The deletion of lemmas is therefore a prerequisite for making room for newly established ones (cf. Rundell, 2015, p. 302). The editors must also face the challenges of turning a former print dictionary, that has been revised to meet the demands of the online versions, back (or maybe forward) to a classical print dictionary again.

Regardless of its form of publication, SAOL is a synchronic dictionary and not a mausoleum for words that are not in use. Considering future studies of the Swedish vocabulary, the list of headwords in SAOL should represent contemporary Swedish.

For users interested in words that are no longer in use, the online resource SAOLhist provides the lists of headwords for all previous editions of the glossary (see section 4 for examples).

3. Methods and Selected Corpora

Kilgarriff (2013, p. 74) states that "building a headword list for a new dictionary (or revising one for an existing title) has never been an exact science" (see also Trap-Jensen, 2013, on lemma selection in relation to corpus content). However, from the SAOL perspective, it is important that the editorial work with both the inclusion of new lemmas and the exclusion of existing ones is scientifically based, as far as possible.

The current SAOL editors have access to both language technology-based methods as well as considerably larger corpora with a broader coverage of different texts, than was the case for previous SAOL lexicographers (see section 2.2). The annotated corpora are available through Språkbanken Text's word research platform Korp (Borin et al., 2012; see also Holmer et al., 2024, pp. 71–72). The corpora consist of newspaper texts, from about 2010 to 2021, from different parts of Sweden. In addition, it contains texts published by Swedish Television (SVT), the public service television company, from the same period, the Swedish version of Wikipedia, and texts from a poetry site with user-generated content. Språkbanken Text's text analysis tool Sparv (Hammarstedt et al., 2022) provides frequency lists of the lemma forms of all words occurring in the corpora. This makes it straightforward to look up how often a headword in the dictionary is used and to produce a list of exclusion candidates. The total number of tokens in the corpora is nearly 3 billion.

4. Results of the Study

Even though several thousand lemmas have been removed between the editions of SAOL, studies have shown that earlier editions of SAOL include relatively many words that are not used in modern texts (Gellerstam, 2005; Berg et al., 2010; see section 2.2).

A preliminary comparison between the lemma list in SAOL 14 and the words that appear in our selected corpora shows that approx. 121,000 of the 126,000 lemmas appear in at least one of their inflectional forms in the text materials.

However, the remaining lemmas, about 4,100, are not used in the texts. This corresponds to approximately 3.5% of the lemmas. The result can be compared with Berg et al. (2010) who show that approximately 10% of the lemmas in SAOL 13 were not used in their reference corpus. Still, it is important to remember that the corpus used in our study is more than 10 times larger, and that the text materials in our corpora represent different genres etc.

Figure 2 shows a selection of examples of lemmas which are not used in the current corpus in any inflectional form.

atomtekniker ('atomic engineer')	rövarromantik ('robber romance')
bofinksbo ('chaffinch nest')	sandalrem ('sandal strap')
dambekantskap ('lady acquaintance')	senvårdag ('late day of spring')
dolikocefali ('dolichocephaly')	sexmejla ('(to send) sex e-mail')
civiltjänstgörare ('person doing civil service')	sladdrare ('gossiper')
eau-de-vie-butelj ('bottle of eau-de-vie')	tevetelefon ('TV telephone')
ellipsformig ('formed like an ellipsis')	tändsticksetikett ('match(box) label')
fingerlik ('fingerlike')	vattenösa ('water scoop')
foderpotatis ('feeding potato')	väsensfrändskap ('being's kinship')
förtursbehandla ('treat with priority')	yawlriggad ('yawl rigged')
förhandlingspraxis ('negotiation practices')	ångerköpthet ('remorse')
grodmansutrustning ('frogman equipment')	änkeleken ('the widow game')
kakelugngaller ('tiled stove grate')	ögonflörta ('(to) flirt with the eyes')

Fig 2: Examples of lemmas in SAOL 14 (2015) which are not used in any inflectional form in the corpus

Among the lemmas in Figure 2, there are nouns such as *dambekantskap* ('lady acquaintance'), *dolikocefali* ('dolichocephaly'), *foderpotatis* ('feeding potato'), adjectives such as *ellipsformig* ('formed like an ellipsis'), *fingerlik* ('fingerlike'), and verbs such as *förtursbehandla* ('treat with priority') and *ögonflörta* ('(to) flirt with the eyes'). Some of the words are registered in several SAOL editions, for example, *sladdrare* ('gossiper'), which has been included since the first edition from 1874. However, many lemmas, such as *fingerlik* ('finger like'), *rövarromantik* ('robber romance'), *tändsticksetikett* ('match(box) label') and *väsensfrändskap* ('being's kinship'), have been present in SAOL since the ninth edition (from 1950), which holds the largest list of lemmas to the present day (approx. 155,000 lemmas).

A more qualitative review of the entire list of so-called non-matches, i.e., SAOL lemmas that are not used in the current corpora, shows that these represent slightly different types of words. Many non-matches consist either of older designations that have been replaced by new ones (such as *hörselvårdsassistent* lit. 'hearing care assistant') or designations for objects, phenomena, activities, etc. no longer in use in the surrounding society (such as *tevetelefon* 'TV telephone' and *skråtobak* 'chewing tobacco').

Furthermore, in the list of non-matches, there are a few, simplex words (such as *pyngel* 'trifles of little value') and a considerable number of solid compounds (such as *rövarromantik* and *tevetelefon*). In particular, complex solid compounds like *grodmansutrustning* ('frogman equipment') and *kakelugngaller* ('tiled stove grate') are frequent in the list.

According to Berg & Malmgren (2010), a lot of "intuitively slightly odd derivations" were excluded in the revisional process of SAOL 14. Despite this, there are still many derivatives among the non-matches. Among them are regularly formed nouns with the derivational suffix *-het* ('-ness'), e.g., *skvalleraktighet* (lit. 'gossipness'), *skäggighet* (lit. 'beardiness'), *snubblighet* (lit. 'stumblyness') and *överklokhhet* (lit. 'over-wiseness'). Nevertheless, these words are extremely rare in modern Swedish texts.

Written Swedish is becoming more gender neutral, and consequently, many derivatives with the suffixes *-inna* and *-erska*, forming older feminine designations, have already been deleted from previous revisions of SAOL (see Gellerstam, 2009b). However, there are still a number of such derivatives among the non-matches, such as *givarinna* ('female donor') and *giftblanderska* ('female poison mixer').

The use of other lemmas in the list, such as *eskimåtröja* ('Eskimo sweater') and *fruntimmerssyssla* (lit. 'women's occupation'), has probably also declined because they are considered less appropriate in modern society, for various reasons. Another reason is that they may have acquired negative connotations (see Gellerstam, 2009b, pp. 82; Petersson & Sköldberg, 2020).

Finally, even though SAOL aims at reflecting the general Swedish language, a lot of specialized words occur in the lemma list, e.g., *dolicocefali* ('dolichocephaly') (in Figure 2). The fact that these words are non-matches in a corpus that even includes Wikipedia might indicate that those lemmas should not be a part of the SAOL lemma list either.

To sum up, it would appear that many of the non-matches could be excluded from the lemma list in SAOL without further ado, but some of the words among the non-matches may require further lexicographic investigation. The word *civiltjänstgörare* ('person doing civil service') is more common in Finnish-Swedish texts, and SAOL has also included a number of Finnish-Swedish words in recent editions (see Gellerstam, 2009b, pp. 76–77). A word like *bofinksbo* may have been included originally to illustrate that compounds with *bofink-* use *s* as a linking morpheme (see Berg et al., 2010, pp. 569–570).

It should also be noted that among the non-matches, there are Swedish words that are probably in use, but mainly in other contexts than those covered by the current corpora (e.g., *krassefrö* 'cress seed' and *vitlökskapsel* 'garlic capsule').

The use of these, as well as the other lemmas in the list of non-hits, will undergo further lexicographic investigations before a possible exclusion of obsolete words takes place for the forthcoming edition of SAOL.

If a lemma is deleted from an edition of the glossary, it will most certainly be searchable in the historical resource SAOLhist in the same way as previously excluded lemmas. A search in SAOLhist with, for example, the search string "%erska" shows several older feminine designations included in previous editions, among them *sanningssökerska* (lit. 'truth seekeress'), *kaffekokerska* (lit. 'coffee cookess') and *konstmakerska* (lit. 'art makeress'). Another search in the resource reveals that excluded SAOL lemmas can be included again. For example, the noun *aeronaut* ('aeronaut') and the adjective *aeronautisk* ('aeronautical') were deleted from the editions 11 (1986) and 12 (1998), but were included again in edition 13 (2006), probably due to their frequency in use.

5. Final Remarks

The revisional work for the 15th edition of SAOL is underway. The study presented here shows that approximately 3.5% of the headwords of the current edition, SAOL 14 (2015), do not occur in any inflectional form in our corpus of 3 billion words. The candidates for exclusions, therefore, amount to approximately 4,000 headwords.

Based on current corpora and modern language technology-based methods, the lexicographical decisions made by the SAOL lexicographers of today can be more efficient and built on a more solid scientific ground than was the case for previous editors.

SAOL needs to be systematically updated to maintain its role as a useful synchronic dictionary relevant for both linguists and non-linguists. For those users who are interested in the Swedish vocabulary of older times, other sources are available, e.g., SAOLhist.

The glossary only includes a selection of contemporary Swedish words, and this selection must be as representative as possible. It should also be noted that all versions of the same edition of SAOL (both print and digital) are expected to list the same headwords. Hence, lemma exclusion is an important building block for this purpose and a necessary complement to lemma inclusion.

Keeping the lemma list in SAOL up to date is also a prerequisite for new areas of use. SAOL could, among other things, strengthen its role as a useful resource in Swedish language technology.

References

- Berg, S., Holmer, L., & Sköldberg, E. (2010). Time to say goodbye? On the exclusion of solid compounds from the Swedish Academy Glossary (SAOL). In A. Dykstra, & T. Schoonheim (Eds.), *Proceedings of the 14th EURALEX International Congress* (pp. 567–576). Fryske Akademy.
- Berg, S., & Malmgren, S.-G. (2010). Varför travintresse ska ut och fotbollsintresse in. *Språkbruk* 2010(4). Retrieved July 30, 2024, from <https://sprakbruk.fi/artiklar/varfor-travintresse-ska-ut-och-fotbollsintresse-in/>
- Borin, L., Forsberg, M., & Roxendal, J. (2012). Korp – the corpus infrastructure of Språkbanken. In *Proceedings of LREC 2012* (pp. 474–478). ELRA.
- Diamond, G. (2016). Making Decisions about Inclusion and Exclusion. In P. Durkin (Ed.), *The Oxford Handbook of Lexicography* (pp. 532–545). Oxford University Press.
- Gellerstam, M. (2005). Ord som går i pension. Om arbetet med att rensa bort gamla ord i SAOL. In R. Vatvedt Fjeld, & D. Worren (Eds.), *Nordiske studiar i leksikografi* 7 (pp. 133–139). Nordisk foreining for leksikografi.

Gellerstam, M. (2009a). Ordförrådets förändring i SAOL . In M. Gellerstam (Ed.), *SAOL och tidens flykt. Några nedslag i ordlistans historia* (pp. 85–104). Norstedts.

Gellerstam, M. (2009b). SAOL i många upplagor. In M. Gellerstam (Ed.), *SAOL och tidens flykt. Några nedslag i ordlistans historia* (pp. 53–83). Norstedts.

ten Hacken, P. (2020). Norms, new words, and empirical reality. *International Journal of Lexicography*, 33(2), 135–149.

Hammarstedt, M., Schumacher, A. Borin, L., & Forsberg, M. (2022). *Sparv 5 User Manual. Research Reports from the Department of Swedish, Multilingualism, Language Technology*. Göteborg: University of Gothenburg 2022/02. Retrieved July 30, 2024, from <https://gupea.ub.gu.se/bitstream/handle/2077/73604/GU-ISS-2022-02%20Sparv%205%20User%20manual.pdf>

Hanks, P. (2019). Lexicography. In R. Mitkov (Ed.), *The Oxford Handbook of Computational Linguistics*, 2 ed. (pp. 473–493). Oxford University Press.

Holmer, L., Hult, A.-K., & Sköldberg, E. (2015). Spell-checking on the fly? On the use of a Swedish dictionary app. In I. Kosem et al. (Eds.), *Electronic lexicography in the 21st century: linking lexical data in the digital age. Proceedings of the eLex 2015 conference* (pp. 356–371). Trojina, Institute for Applied Slovene Studies/Lexical Computing Ltd.

Holmer, L., Lillieström, A., Sköldberg, E., & Uppström, J. (2024). SAOL och svensk språkvetenskaplig infrastruktur – nu och i framtiden. In E. Volodina et al. (Eds.), *Proceedings of the Huminfra Conference (HiC 2024)* (pp. 68–75). LiU Electronic Press.

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, Faculty of Arts.

Petersson, S., & Sköldberg, E. (2020). To discriminate between discrimination and inclusion: A lexicographer's dilemma. In Z. Gavriilidou, M. Mitsiaki, & A. Fliatouras (Eds.), *Proceedings of the 19th EURALEX International Congress, Vol. 1*. (pp. 381–386). Democritus University of Thrace.

Rundell, M. (2015). From Print to Digital: Implications for Dictionary Policy and Lexicographic Conventions. *Lexikos*, 25, 301–322.

Sánchez Ibáñez, M., & Maroto, N. (2021). Beyond timelines: the challenges of combining theoretical premises and speakers' insights about the assessment, validation and inclusion of Spanish neologisms in dictionaries. *International Journal of Lexicography*, 34(3), 358–381.

SAOL 11: *Svenska Akademiens ordlista*. (1986). 11th edition. Norstedts.

SAOL 12: *Svenska Akademiens ordlista*. (1998). 12th edition. Norstedts.

SAOL 13: *Svenska Akademiens ordlista*. (2006). 13th edition. Norstedts.

SAOL 14: *Svenska Akademiens ordlista* (2015). 14th edition. Norstedts. Retrieved May 7, 2024, from <https://svenska.se/saol/>

SAOLhist: Välkommen till SAOLhist. Retrieved May 7, 2024, from <https://spraakbanken.gu.se/saolhist/>

Sköldberg, E., Holmer, L., Volodina, E., & Ildikó Pilán (2019). State-of-the-art on monolingual lexicography for Sweden. *Slovenščina 2.0*, 7(1), 13–24.

Svensson, A. (2019). Här är orden som försvinner. *Språktidningen* 2019(8). Retrieved May 7, 2024, from <https://spraktidningen.se/artiklar/har-ar-orden-som-forsvinner/>

Trap-Jensen, L. (2013). Researching lexicographical practice. In H. Jackson (Ed.), *Bloomsbury Companion to Lexicography* (pp. 19–30). Bloomsbury.

Vikør, L. S. (2009). Lexicography and language planning in Scandinavia and the Netherlands. In S. Nielsen, & S. Tarp (Eds.), *Lexicography in the 21st Century. In honour of Henning Bergenholtz* (pp. 123–143). John Benjamins.

Acknowledgements

This work has been funded by the Swedish Academy, and by Nationella språkbanken – jointly funded by its 10 partner institutions and the Swedish Research Council (2018–2024; dnr 2017-00626).

Contact information

Louise Holmer

University of Gothenburg, Department of Swedish, Multilingualism, Language Technology
louise.holmer@svenska.gu.se

Ann Lillieström

University of Gothenburg, Department of Swedish, Multilingualism, Language Technology
ann.lilliestrom@gu.se

Emma Sköldberg

University of Gothenburg, Department of Swedish, Multilingualism, Language Technology
emma.skoldberg@svenska.gu.se

Jonatan Uppström

University of Gothenburg, Department of Swedish, Multilingualism, Language Technology
jonatan.uppstrom@svenska.gu.se