Revealing Semantic Variation in Swedish Using Computational Models of Semantic Proximity – Results From Lexicographical Experiments

Sköldberg, Emma; Virk, Shafqat Mumtaz; Sander, Pauline; Hengchen, Simon; Schlechtweg, Dominik

Emma Sköldberg, Shafqat Mumtaz Virk, Pauline Sander,
Simon Hengchen, and Dominik Schlechtweg

REVEALING SEMANTIC VARIATION IN SWEDISH USING COMPUTATIONAL MODELS OF SEMANTIC PROXIMITY

Results From Lexicographical Experiments

Abstract The paper reports a pilot study on the detection of lexical semantic variation in modern Swedish. The starting point of the study is the meaning descriptions of around 65,000 headwords in 'The Contemporary Dictionary of the Swedish Academy' (SO, 2021) covering approximately 100,000 different senses. In our work, we aim to explore the potential of the latest computational methods to discover outdated definitions in SO and update them. For this, we make use of the DURel tool (Schlechtweg et al., 2018, 2024) which relies on state-of-the-art language models for the automatic semantic analysis of word usages. The work resulted in drawing lexicographers' attention to both main senses and subsenses that should be added to the dictionary. It has also demonstrated that certain meaning descriptions in SO are too general and should be split in accordance with the current principles for the semantic descriptions in the dictionary.

Keywords lexicography; language technology; lexical semantics; semantic variation and change; automatic semantic analysis; word sense induction, semantic proximity; Swedish; DURel

1. Introduction

As shown by e.g., Lau et al. (2012), the meaning descriptions in dictionaries frequently become outdated, and hence require continuous updates. In order to meet the needs in practical lexicography in relation to this area, Cook et al. (2013), Nimb et al. (2020) among others, have conducted lexicographically-oriented studies on the automatic detection of semantic variation and change. However, the research in this direction, including method development could be taken much further, especially in relation to relatively under-resourced languages such as Swedish (cf. Cavallin, 2012).

In this paper, we report on ongoing research regarding the automatic detection of lexical semantic variation and change in modern Swedish, and how this work benefits Swedish lexicography. The starting point of the experiments in the study is the meaning descriptions in *Svensk ordbok utgiven av Svenska Akademien* ('The Contemporary Dictionary of the Swedish Academy' (SO, 2021) (see e.g., https://svenska.se/), which constitutes the most complete description of the Swedish vocabulary of today (see Section 2).

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

In this work, we aim to explore the potential of the latest computational methods to discover outdated definitions in SO and update them. For this, we make use of the annotation tool *DURel* (Schlechtweg et al., 2018, 2024) which relies on state-of-the-art language models for automatic semantic analysis of word usages (see Section 3). In more concrete terms, we have conducted two experiments. The first experiment aimed to validate the feasibility and prove the usefulness of DURel and tune hyper-parameters of its underlying computational language models. The second experiment aimed, by using DURel, to revise semantic descriptions of headwords already included in SO.

The outline of the paper is as follows: In Section 2 we present SO with focus on the meaning descriptions in the dictionary. Section 3 includes a brief overview of the annotation tool DURel as well as the corpus that we have been using in the experiments. After that, in Section 4, we present the two experiments that we have conducted and finally, in Section 5, we summarize and discuss our results.

2. SO and Its Sense Descriptions

As already mentioned, the point of departure in our research is SO, a definition dictionary with about 65,000 headwords describing the vocabulary of modern Swedish. SO, which is corpus-based, covers approximately 100,000 different senses and it is primarily aimed at users with Swedish as their mother tongue, but also advanced language learners of Swedish.

SO is a subset of a very extensive lexical database, which has been under continual development at the University of Gothenburg since the 1970s (see Ralph et al., 1977). For each edition, SO will provide as complete information as possible on recurrent words and expressions in Swedish general language. This information includes the words' spelling, pronunciation, inflection, etymology etc. However, the emphasis in SO is on the meanings and uses of the words (see e.g., Malmgren & Sköldberg, 2013).

As is well-known, practical lexicographic work is tradition-bound and all the editions of SO follow certain principles described in e.g., Ralph et al. (1977) and Järborg (1989). These principles also apply to the meaning descriptions in the dictionary. According to Svensén (2009, pp. 211–212), the polysemy structure of the words in a dictionary can be described 1) linearly, i.e., as a number of discrete units arranged in a sequence or 2) hierarchically, as a number of main/core senses, to which groups of subsenses/shades are associated. In SO, the latter approach is adopted. Furthermore, in SO, the relationship between every main sense and the subsense(s) is explicitly specified (in terms of meaning extension, meaning specialisation, metaphorical (figurative) use etc.) To illustrate how the two different polysemy structure principles work, consider the example *aptitretare* (appetizer) with one more literal and one more figurative sense. When the polysemy structure is described linearly in the dictionary, these two meanings are listed as 'meaning 1' and 'meaning 2' and when the polysemy structure is hierarchical, the two meanings are listed as 'meaning 1' and 'meaning 1a', because the second meaning is considered to be a figurative subsense of the main sense.

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

As Lew (2013, p. 287), among others, have shown, two opposing strategies, known as *lumping* and *splitting* have been identified when it comes to specifying senses in monolingual dictionaries. The first strategy aims to minimize the number of senses so that each of them cover as much semantic ground as possible. The second strategy tends to result in a rather larger number of finely distinguished senses. The lexicographers' choice of strategy often depends on the current type of dictionary and the dictionary's target users (Atkins & Rundell, 2008, pp. 267–268). The semantic description in a dictionary that primarily addresses learners at beginner level is frequently characterized by lumping while the semantic description in a dictionary that addresses linguists, etc. is characterized by splitting. In SO, the analysis and the distinction into different senses is relatively "fine-grained". In other words, the SO lexicographers are rather splitters than lumpers.

Since the semantics of the headwords in SO is fundamental, an important part of the revision work between the editions is to examine if the meanings of the headwords have changed – and in that case, how they have changed. However, this work is very time-consuming. As pointed out by Petersson & Sköldberg (2021), the SO lexicographers currently do not use any computational methods for discovering semantic change in a systematic way.

3. Tool and Corpora

In this section, we shortly present the tool and the corpus that we have used in the experiments.

The DURel annotation tool (https://durel.ims.uni-stuttgart.de/) provides functionalities to upload sets of word usages for a particular word of interest and to annotate the semantic proximity (a.k.a. similarity/relatedness) between these usages either with humans or computers. This information (usages with their proximities) can then be visualized in a graph and clustered into sets of high-proximity usages, which can be interpreted as word senses.

The tool allows to inspect the graph and to adjust multiple annotation, cluster and visualization parameters. In this way, we can easily infer lexicographically relevant semantic information such as the inferred number of senses or their changes over time. Find an example of a DURel graph concerning the usage of the noun *baksida* ('back, downside, disadvantage') in the SVT corpus in Figure 1.

DURel offers multiple computational annotators for semantic proximity based on state-of-the-art pre-trained Word-in-Context models (Pilehvar & Camacho-Collados, 2019). We rely on the most recent model, XL-LEXEME, which was optimized for Lexical Semantic Change Detection (Cassotti et al., 2023). This model maps each word usage onto a contextualized distributional-semantic vector representation and then uses cosine similarity to estimate the semantic proximity between two usage vectors.

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

For our experiments, we have used the SVT corpus which consists of texts published by the Swedish public service television company between 2004 and 2021. The corpus covers about 200 million tokens and is available through Korp, Språkbanken's word research platform (Borin et al., 2012).

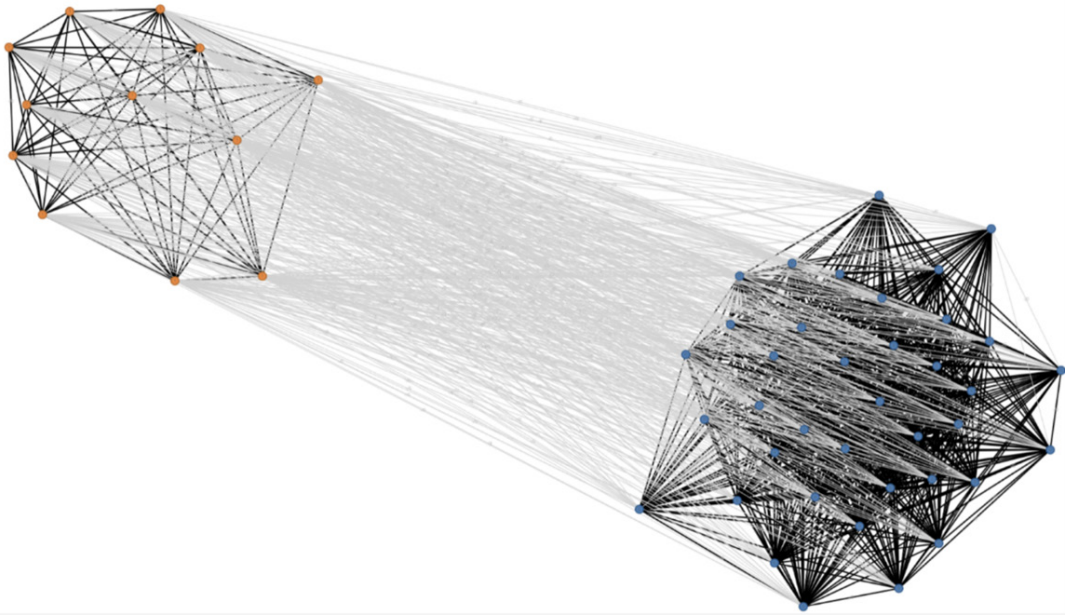

Fig. 1: The noun *baksida* ('back, downside, disadvantage'), based on corpus samples from the SVT corpus, forming two semantic clusters in DURel.

Figure 1 visualizes usages of the word *baksida* as nodes while edge weights are given by the cosine similarity between usages annotated with the XL-LEXEME model. The graph was clustered using Correlation Clustering (Schlechtweg et al., 2020). DURel highlights two semantic clusters, one with blue dots and another with orange dots among the uses of the word *baksida* in the corpus. If the users click on individual dots in the DURel front-end interface, they will find the corresponding corpus sample. In this case, among the blue dots there are samples including word combinations such as "polishusets baksida" ('the back of the police house') and "baksidan av låret" ('the back of the thigh'). Among the orange dots, there are samples such as "framgångens baksida" ('lit. the back side of the success')" and "baksidan med droger" (lit. 'the back side with drugs'). In other words, DURel has here succeeded in pointing out at least two different senses of the word in the corpus, the latter of which is clearly metaphorical.

4. The Experiments

Petersson & Sköldberg (2021) show, by discussing different types of examples in Swedish, that the phenomenon of lexical semantic variation and change, is multifaceted and diverse. Hence, within our research, we strive to approach the phenomenon from partially different angles. One way is to start from SO dictionary articles concerning headwords with more than one sense and evaluate whether DURel manages to cluster these senses among the uses of the word in a corpus (see experiment 1). Another way

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

is to start with headwords in SO with only one sense and see if DURel predicts two or more sense clusters of the word in a corpus. Two or more clusters may indicate that the word has more than one sense in the texts and that the semantic description in SO should be revised (see experiment 2).

4.1 Experiment 1

In experiment 1, we started by selecting a set of 25 SO-headwords, based on their semantic characteristics: all of them have a main sense and one or more subsenses. An example is the adjective *enkelspårig* with the main sense 'one-track' and the figurative subsense 'simplistic, superficial, narrow-minded'.

Next, 50 random usages (sentences) of each headword were extracted from the SVT corpus. After that, SO lexicographers assessed all the sentences and classified if the meaning of a certain word usage was covered in the dictionary and, if that was the case, which sense it was (the main sense, a subsense etc.). The goal of this work was partly to get an indication of the semantic quality of the current dictionary articles, partly to create a so-called manually curated gold standard (see below). This closer examination of the corpus samples showed, among other things, that the semantic description of 7 of the selected headwords in the dictionary were not comprehensive enough and should be supplemented. For example, in 12 out of 50 occurrences in the corpus, the verb *explodera* ('explode') meant 'increase significantly' (see e.g., "Försäljningen av miljöbilar har exploderat det senaste året" ('Sales of green cars have exploded in the past year')). In the sample you find a well-established figurative use of the word *explodera* that is currently missing in SO and according to the editorial principles of the dictionary (see Ralph et al., 1977, and Järborg, 1989), that sense should form a subsense of its own in the dictionary article. Consequently, *explodera* and 6 other headwords were, based on quality reasons, excluded from the experiment. Even if the quality of SO is high in general, this result supports previous conclusions by e.g., Blensenius et al. (2021) concerning the development potential of the semantics in the dictionary.

After that, the remaining 18 headwords were uploaded to DURel. Each of the 50 usages of the words was paired with every other usage producing a total of 2450 usage pairs for each headword. These pairs were then annotated automatically for sense similarity (i.e., if a given headword is used in the same sense in the usage pair or not) using XL-LEXEME (Cassotti et al., 2023), a state-of-the-art language model for lexical semantic change detection. The annotations were clustered with the correlation clustering algorithm described in Schlechtweg et al. (2020). The clustering algorithm has a threshold hyper-parameter that was tuned to yield optimal results on the gold data (see below). After uploading the usages, all further steps were performed using the DURel front-end interface. Figure 2 shows the two resulting clusters for *enkelspårig*, one with blue dots and another with orange dots, corresponding to the literal and figurative senses of this headword in SO.

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

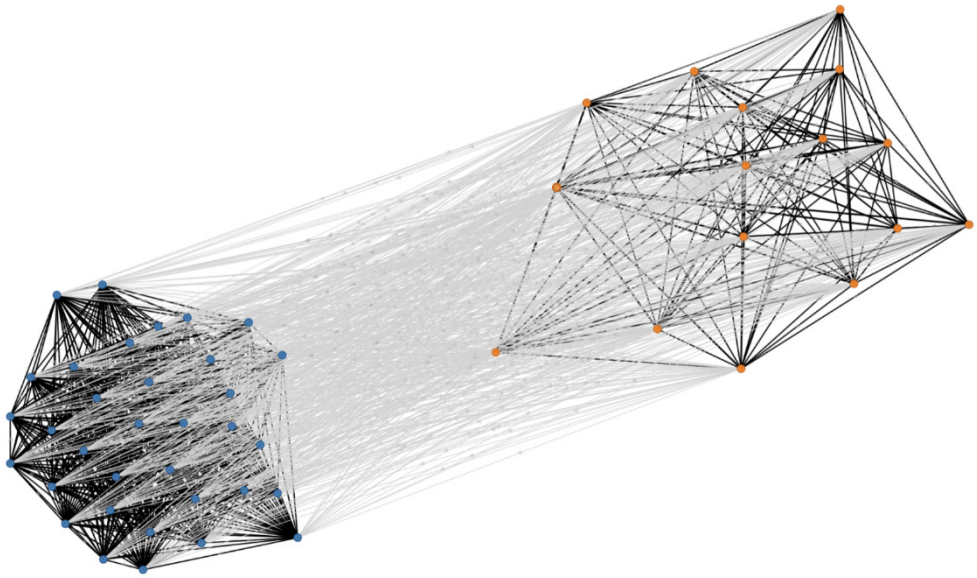

Fig. 2: DURel sense cluster for the headword *enkelspårig* ('one-track; 'simplistic, superficial, narrow-minded').

In the front-end interface, you see, by clicking on the blue spots, corpus samples like "Den enkelspåriga järnvägen mellan Motala och Hallsberg är idag en flaskhals ..." ('The one-track railway between Motala and Hallsberg is a bottleneck today ...'). Hence, in this particular context, the literal, main sense of the adjective is used. By clicking on orange dots, you see uses like "De tror att vi är enkelspåriga lantisar, de tror att vi är trångsynta, att vi är rasister och homofober." ('They think we're narrow-minded peasants, they think we're bigoted, that we're racists and homophobes.'). Here, the figurative sense of the word, which corresponds to the subsense in SO, is relevant. To conclude, in this case, DURel has managed to distinguish between the two senses of *enkelspårig*.

In the next step of the experiment, the gold clusters, i.e., the results of the manual analysis of the language samples by the lexicographers, were evaluated against the automatically generated clusters by DURel. Table 1 shows the Adjusted Rand Index (ARI; Hubert & Arabie, 1985) values, which is a measure of the similarities between clusters, here the manually curated gold clusters and the clusters formed by DURel. We clustered all graphs with different thresholds in the range of (0.3 0.325 0.35 ... 0.7) and selected the threshold with highest ARI for further analysis and experiments, resulting in a threshold of 0.6 for clustering.

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

Table 1: Cluster evaluation based on ARI

Headword	**ARI**	**Headword**	**ARI**
enkelspårig ('one-track, simplistic')	1.0	kriga ('make war')	0.614
fasad ('facade')	1.0	rutten ('rotten')	0.299
ofantlig ('immense')	1.0	ventilera ('ventilate')	0.291
baksida ('back')	0.863	lirka ('tinker, coax')	0.251
fotavtryck ('footprint')	0.84	vansinnig ('insane')	0.237
klimat ('climate')	0.772	hagla ('fall hail')	0.228
bagage ('baggage')	0.758	skör ('fragile')	0.068
vissen ('withered')	0.645	hemmaplan ('setting')	0.0
tvärnita ('jam on the breaks')	0.642	kapitulera ('capitulate')	-0.008
Average	0.528		

ARI generally ranges from -0.5 to 1. Values closer to 1 signify better agreement between the gold clusters and the derived clusters. A value of 0 suggests a random clustering while negative values indicate dissimilarity.

As can be noted, words such as the already mentioned *baksida* ('back, downside, drawback') and *enkelspårig* ('one-track, simplistic, narrow-minded') but also the noun *fasad* ('facade'), the adjective *ofantlig* ('immense'), and the noun *fotavtryck* ('footprint') have relatively high ARI values, indicating stronger alignment between their respective gold clusters and derived clusters. In contrast, the adjective *vansinnig* ('insane') and the verbs *hagla* ('hail') and *kapitulera* ('capitulate') show ARI values of 0 or slightly negative. An average of 0.528 across all words represent a reasonable overall similarity, and the results suggests that the automatically derived clusters generally encode meaningful semantic information which is useful for lexicographers.

In addition to the automatic ARI evaluation, the DURel clusters were analyzed qualitatively by the lexicographers, which provided useful insights. For example, an examination of the figurative examples of *bagage* ('luggage') revealed that some of the uses constitute different variants of the idiomatic expression, *ha något i bagaget* ('to have something in the luggage') (see Schlechtweg et al., 2024). The idiom is not covered in the current edition of SO and will be included in the next edition.

Another example is the verb *hagla* (lit. 'fall hail'). In SO, the headword has two senses, the main sense 'fall hail' and the figurative subsense 'appear (over someone) in large quantity // mostly about abstract phenomena'. As shown in Figure 3, DURel predicted three main semantic clusters which might be an indication that the word has three senses.

Among the blue-marked uses in Figure 3, the main weather sense of the verb in SO is represented. However, among the uses marked with an orange dot, there are sentences like "De [ungdomarna] står ofta på avstånd och kastar och för polisen blir det då att springa mot dem medan stenar haglar." ('They [the young people] often

stand at a distance and throw, and the police then have to run towards them while stones are hailing'.) Furthermore, among the green-marked dots there are samples like "Smädeorden haglade" ('the vituperations rained/came thick and fast'). In this case the SO lexicographers also concluded that, based on the editorial principles, the current subsense in the entry *hagla* is too wide. It should be split into two subsenses, one based on meaning extension concerning more concrete objects (like stones etc.), and the other one figurative, concerning abstract phenomena like vituperations. In other words, even if the correspondence between the gold clusters and the derived clusters is not perfect (as in the case of *hagla*; see Table 1 above), the lexicographers can get useful information about the meanings of a headword by using DURel.

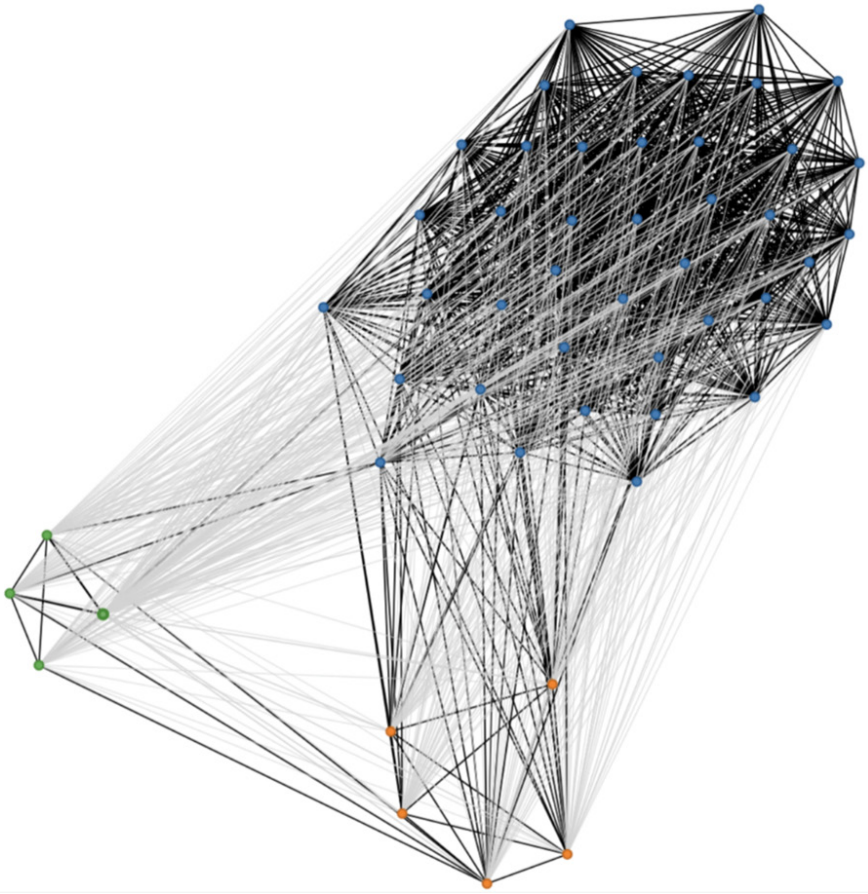

Fig. 3: DURel sense cluster for the verb *hagla* ('fall hail; appear (over someone) in large quantity // mostly about abstract phenomena')

4.2 Experiment 2

After getting promising results from experiment 1, we started the second experiment with the major aim mentioned above. The methodology was the same, but this time we focused on the number of senses recorded in SO and compared that number with the number of clusters inferred for a given headword in the corpus using DURel. A discrepancy in the number of senses and the recorded clusters can be an indicator that the dictionary article is outdated. Such headwords can be prioritized in order to update the dictionary.

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

During the first phase, we concentrated on the 39,000 nouns, adjectives and verbs with only one sense recorded in SO. A cursory review of this extensive data by the lexicographers showed that many headwords probably have more senses than the ones listed in the dictionary. An example is the noun *botemedel* which according to SO means "medel för att undanröja ett problem // ofta sjukdom men även allmännare" ('means to eliminate a problem // often illness but also more generally'). The headword *botemedel* should, in line with current principles for meaning descriptions in the dictionary, be split into one main sense and one subsense (since the word is also used in more general contexts) (see Ralph et al., 1977; Järborg, 1989). The medical-related main sense could be illustrated by corpus-based examples including expressions like "jakten på ett botemedel mot AIDS" ('the search for a cure for AIDS') and the more general subsense by "Finns det ett botemedel mot trångsynthet och idioti?" ('Is there a cure for bigotry and idiocy?'). A search in the SO database shows that there are about 180 other headwords in the dictionary with senses that should be split in the same way. In other words, based on individual words in the list with monosemic headwords, the lexicographers have found groups of headwords in SO whose semantic description could be improved.

Since DURel has not been previously tested at that large scale, for a quick proof of concept we selected a set of random 281 headwords (nouns, adjective and verbs) in SO with only one sense registered. All headwords have at least 25 occurrences in the SVT corpus, which is an important prerequisite for being able to notice semantic patterns of the words.

Next, usages were extracted from the corpus and clusters were inferred as explained previously (see section 3) using the threshold parameter tuned in experiment 1. It appeared that 215 out of 281 words were predicted to have only one cluster (and hence one sense). Among those headwords you find nouns such as *brevinkast* ('letterbox') and *tisdag* ('Tuesday'), adjectives like *oåterkallelig* ('irreversible') and *tunisisk* ('Tunisian') and verbs like *orsaka* ('cause') and *sukta* ('long (in vain)'). According to the SO lexicographers' assessment, these headwords, like many others in the list, have only one sense. The semantic description of those headwords in SO is thus satisfactory.

However, 49 of the 281 headwords in the experiment have two clusters. Furthermore, 9 have three clusters, 6 have four clusters, 1 has five clusters, and 1 has six clusters in DURel. In other words, according to the annotation tool, these headwords were predicted to have between 2 to 6 senses.

While a full-scale proper evaluation of the results in this experiment still needs to be carried out, we were able to find promising cases by looking at some of the examples with more than one cluster. One example is the noun *lydnad* ('obedience') for which DURel detected two clusters, colored blue and orange in Figure 4.

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

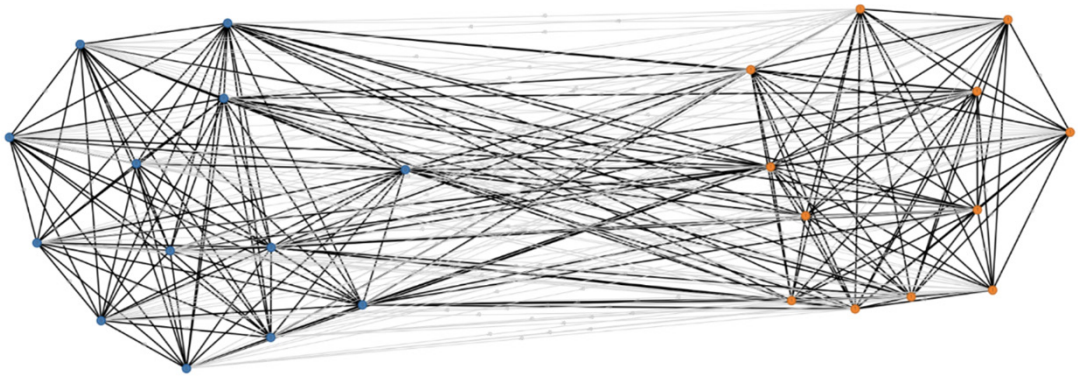

Fig. 4: DURel sense cluster for the noun *lydnad* ('obedience').

The sense cluster by DURel indicates that the noun *lydnad* has at least two different senses which was confirmed by the lexicographers who inspected the clusters and the corpus samples manually. They found that the cluster made up of orange dots includes corpus samples such as "Nunnor följer de tre klosterlöftena: att leva i celibat, fattigdom och lydnad. ('Nuns follow the three monastic vows: to live in celibacy, poverty and obedience.') This example illustrates the meaning already recorded in SO. The cluster with blue dots, on the other hand, includes samples with word combinations like "tävla i lydnad" ('compete in obedience') and "nordiska mästerskap i lydnad" ('Nordic championships in obedience'). They represent a specialization of the sense which is related to dog- and equestrian sports. This specific meaning of the noun is not recorded in the current edition of SO and will be added in the next update.

Among the headwords with only one sense in SO but with two or more confirmed semantic clusters in the analysis by DURel one may also find the noun *slutspurt* ('final spurt' but also 'finish'), *brottningsmatch* ('wrestling-match') and *avstickare* ('detour' but also 'digression'). Through the tool, the SO lexicographers have, among other things, been drawn to the figurative meanings of the three headwords in the corpus, which will result in improved dictionary articles.

Another example is the noun *nedtagning* with no less than five automatically generated semantic clusters by DURel. However, according to SO, the word has just one sense: 'dampening the movement of a (high) ball with the body (chest) or leg // and placing it in front of the feet; in football'. In this case, it is not a subsense that is missing – the existing one is already very specific – and the lexicographers agree that a main, more basic, sense, which has to do with the act of 'taking down', should be added. The main sense of the word could be illustrated by corpus-based examples like "nedtagning av träd" ('removal of trees') och "uppsättning och nedtagning av belysning" ('setting up and taking down lighting').

However, based on the results of this experiment, it is also clear that the tool can be improved. For example, when it comes to nouns like *fransos* ('Frenchman') and *kulstötare* ('shot-putter'), DURel shows two clusters, but the words have only one sense in SO and in the corpora. Furthermore, the headword *minfält* has just one cluster in DURel, but a

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

closer look at the samples by the SO lexicographers shows that the word has two senses in Swedish: 'minefield' and 'delicate subject'. The figurative meaning of the word is still missing in SO, so the dictionary article should be supplemented.

5. Results and Final Discussion

In this paper we presented two experiments. The aim of the first was to validate the feasibility and prove the usefulness of the annotation tool DURel and its underlying computational language models for semantic variation and change detection. The aim of the second one was to use DURel to revise semantic descriptions of headwords in SO. To sum up, a discrepancy between the number of senses recorded in SO and the number of clusters identified by DURel can be used as an indicator that the description in the dictionary is outdated. Such headwords can be prioritized for manual inspection in order to update the dictionary entry.

The experiments are still small-scaled, but the results are nevertheless promising. So far, the work has, among other things, brought the lexicographers' attention to subsenses (mostly figurative uses of the headwords but also meaning extensions and specializations) that should be added to the dictionary. Also missing main senses have been discovered. The work has also demonstrated that certain meaning descriptions in SO are too general and should be split in accordance with the current principles for the semantic descriptions in the dictionary. So far, we have mainly found well-established senses that should have been included in the latest edition of SO, but we have also found completely new senses in Swedish.

In this study we only used relatively modern Swedish corpora, and hence, we first of all registered lexical semantic variation. One way of capturing lexical semantic changes is by comparing semantic clusters based on corpora from different periods of time. If DURel detects different numbers of clusters for a given word in different corpora, it may indicate that the word's semantics have changed over time.

The study also provided new insights to improve the computational model and DURel. Throughout our study we considerably improved the potential of DURel to perform large-scale computational predictions and to visualize and manually inspect these. The lexicographers' analysis of the clustered graphs can inform the future development of computational models, e.g., by retraining models on the errors identified by lexicographers.

The computer-aided methods are expected to streamline and enhance the lexicographic work on semantic variation and change detection. In the long run, this will help lexicographers to revise SO and other dictionaries in an efficient way.

References

Atkins, B. T. S., & Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. Oxford University Press.

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress*. Institute for the Croatian Language.

Blensenius, K., Sköldberg, E., & Bäckerud, E. (2021). Finding gaps in semantic descriptions. Visualisation of the cross-reference network in a Swedish monolingual dictionary. In I. Kosem, M. Cukr, M. Jakubíček, J. Kallas, S. Krek, S., & C. Tiberius (Eds.), *Electronic lexicography in the 21st century. Proceedings of the eLex 2021 conference. 5–7 July 2021* (pp. 247–258). Online. Brno.

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). Istanbul: ELRA.

Cassotti, P., Siciliani, L., de Gemmis, M., Semeraro, G., & Basile, P. (2023). Xl-lexeme: Wic pretrained model for cross-lingual lexical semantic change. In *Proceedings of the 61th Annual Meeting of the Association for Computational Linguistics*. (Volume 2: Short Papers) (pp. 1577–1583). Online. Association for Computational Linguistics.

Cavallin, K. (2012). Exploring semantic change with lexical sets. In R. V. Fjeld, & J. M. Torjusen (Eds.), *Proceedings of the XV EURALEX international congress* (pp. 1018–1022). Oslo.

Cook, P., Lau, J. H., Rundell, M., McCarthy, D., & Baldwin, T. (2013). A lexicographic appraisal of an automatic approach for detecting new wordsenses. In I. Kosem, J. Kallas, P. Gantar, S. Krek, M. Langemets, & M. Tuulik (Eds.), *Proceedings of eLex 2013 conference* (pp. 49–65). Ljubljana/Tallinn: Trojina, Institute for Applied Slovene Studies/Eesti Keele Instituut.

Hubert, L., & Arabie, P. (1985). Comparing partitions. *Journal of Classification, 2*, 193–218.

Järborg, J. (1989). *Betydelseanalys och betydelsebeskrivning i Lexikalisk databas. Preliminär version.* Göteborg: Göteborgs universitet.

Korp: Språkbanken's word research platform. Retrieved May 15, 2024, from https://spraakbanken.gu.se/korp

Lau, J. H., Cook, P., McCarthy, D., Newman, D., & Baldwin, T. (2012). Word sense induction for novel sense detection. In W. Daelemans (Ed.), *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 591–601). Avignon, France: Association for Computational Linguistics.

Lew, R. (2013). Identifying, ordering and defining senses. In H. Jackson (Ed.), *The Bloomsbury companion to lexicography* (pp. 284–302). Bloomsbury.

Malmgren, S.-G., & Sköldberg, E. (2013). The Lexicography of Swedish and other Scandinavian Languages. *International Journal of Lexicography, 26*(2), 117–134. http://dx.doi.org/10.1093/ijl/ect008

Nimb, S., Hartvig Sørensen, N., & Lorentzen, H. (2020). Updating the dictionary: Semantic change identification based on change in bigrams over time. *Slovenščina* 2.0 8(2), 112–138. https://doi.org/10.4312/slo2.0.2020.2.112-138

Petersson, S., & Sköldberg, E. (2021). Semantic change in Swedish – from a lexicographic perspective. In N. Tahmasebi, L. Borin, A. Jatowt, Y. Xu, & S. Hengchen (Eds.), *Computational approaches to semantic change* (pp. 149–167). Language Science Press.

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.

Ralph, B., Järborg, J., & Allén, S. (1977). *Svensk ordbok och Lexikalisk databas. Förstudierapport.* Göteborg: Göteborgs universitet.

Schlechtweg, D., Schulte im Walde, S., & Eckmann, S. (2018). Diachronic Usage Relatedness (DURel): A framework for the annotation of lexical semantic change. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 169–174). New Orleans, Louisiana.

Schlechtweg, D., McGillivray, B., Hengchen, S., Dubossarsky, H., & Tahmasebi, N. (2020). SemEval-2020 Task 1: Unsupervised Lexical Semantic Change Detection. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation* (pp. 1–23). Barcelona (online), International Committee for Computational Linguistics.

Schlechtweg, D., Virk, S. M., Sander, P., Sköldberg, E., Theuer Linke, L., Zhang, T., Tahmasebi, N., Kuhn, J., & Schulte im Walde, P. (2024). The DURel Annotation Tool: Human and Computational Measurement of Semantic Proximity, Sense Clusters and Semantic Change. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations* (pp. 137–149). St. Julians, Malta: Association for Computational Linguistics.

SO 2021: *Svensk ordbok utgiven av Svenska Akademien* ['The Contemporary Dictionary of the Swedish Academy']. (2021). 2nd edition. Retrieved May 15, 2024, from https://svenska.se/

Svensén, B. (2009). *A handbook of lexicography: The theory and practice of dictionary-making.* Cambridge University Press.

SVT (Sweden Television) corpus. Retrieved May 15, 2024, from Språkbanken's word research platform Korp https://spraakbanken.gu.se/korp

Taher Pilehvar, M. & Camacho-Collados, J. (2019). WiC: the word-in-context dataset for evaluating context-sensitive meaning representations. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.* Volume 1 (Long and Short Papers) (pp. 1267–1273). Minneapolis, Minnesota: Association for Computational Linguistics.

Acknowledgements

This work has jointly been funded by the research program Change is Key! supported by Riksbankens Jubileumsfond (under reference number M21-0021) and the Swedish Academy.

Contact information

Emma Sköldberg
University of Gothenburg, Department of Swedish, multilingualism, language technology
emma.skoldberg@svenska.gu.se

Shafqat Mumtaz Virk
University of Gothenburg, Department of Swedish, multilingualism, language technology
shafqat.virk@svenska.gu.se

Pauline Sander
University of Stuttgart, Institute for Natural Language Processing
pauline.sander@ims.uni-stuttgart.de

Simon Hengchen
iguanodon.ai and Université de Genève
simon@iguanodon.ai

Dominik Schlechtweg
University of Stuttgart, Institute for Natural Language Processing
dominik.schlechtweg@ims.uni-stuttgart.de

This paper is part of the publication: Despot, K. Š., Ostroški Anić, A., & Brač, I. (Eds.). (2024). *Lexicography and Semantics. Proceedings of the XXI EURALEX International Congress.* Institute for the Croatian Language.