Chapter 4

Semantic change in Swedish – from a lexicographic perspective

Stellan Petersson & Emma Sköldberg
University of Gothenburg

In this chapter, we examine semantic change in the general vocabulary of present-day Swedish and its lexicographic description. We discuss the question of whether automatic and semi-automatic methods of computational linguistics are relevant to lexicography and conclude that such methods can facilitate, formalize, and sharpen lexicographic investigations of semantic change.

1 Introduction

Several efforts have been made to automate or semi-automate parts of the process of dictionary compilation, including the building of headword lists and identification of collocations (Cook et al. 2014). Automatic methods for finding linguistic examples have also been developed (see, e.g., Kilgarriff et al. 2008, Pilán 2016). Furthermore, there are computational linguistic studies that examine semantic changes in large text corpora (e.g. Cavallin 2012, Cook et al. 2013, Nimb et al. 2020). A central aim of studies of this kind is to make lexicographic work more efficient; another, related aim is to introduce more systematicity into the process of dictionary construction. The results of studies like these are, of course, relevant to practical dictionary editing. In the ongoing work on the forthcoming second edition of the dictionary *Svensk ordbok utgiven av Svenska Akademien* ('The contemporary dictionary of the Swedish Academy', henceforth "SO"), semantic changes on the lexical level are important. However, the editorial group (of which the authors are members) currently lacks formal, computational methods for discovering semantic changes on the lexical level.

The main purpose of this chapter is to discuss lexicographical problems that are associated with a number of Swedish examples, where each type of example represents an area of research on lexical semantic change. Furthermore, the chapter addresses how computational linguistics and language technology can facilitate lexicographical research in this area.

In Section 2, we begin by providing a general characterisation of our lexicographical framework, focusing on the database from which different editions of SO are based. We then proceed to a discussion of a number of Swedish examples, in Section 3, and explore several lexicographical issues that are relevant to the database and its development. Finally, in Section 4, we turn to the interface between lexicography and computational linguistics and provide some remarks on research in this area relevant to the work on SO.

2 SO and the lexical database

The focus in this chapter is on SO, a definition dictionary containing about 65,000 headwords describing the general vocabulary of modern Swedish. The emphasis in the dictionary is on the meanings and uses of the words. SO, which is corpus-based, is primarily aimed at users with Swedish as their mother tongue, but also at learners with good knowledge of Swedish. The first edition of SO was published in book form in 2009. It is now also available in a digital format, as an app and through the Swedish Academy's dictionary portal, svenska.se. The second edition of SO is scheduled to be published in 2021, but only in digital format.

SO is a subset of a very extensive lexical database (currently including approximately 200,000 headwords) which has been under continual development at the University of Gothenburg (GU) since the 1970s. According to a collaboration agreement between the Swedish Academy and GU from 2010, GU will further develop and maintain the extensive database until 2060. The database, which a research group at the Department of Swedish is responsible for, consists of new words, word forms, and word connections continuously incorporated into the database. The publisher's aim is to release a new edition of SO at least every ten years. For each edition, SO will provide as complete information as possible about every important word and expression in Swedish. This information includes the word's spelling, pronunciation, inflection, style, emotive charge, and meaning. Each word entry will be illustrated with language examples but also with examples of phraseology and constructions.

The 2009 edition of SO has several merits, but because large parts of the dictionary articles were compiled in the 1980s, the dictionary can be improved upon

and modernized in several ways. A fundamental part of the revision work with SO has been to review the headword list. Since the semantic description of the headwords is so central to the dictionary, an important part of the revision work has been to examine whether the meanings of the headwords have changed since the first edition of the dictionary. In conjunction with this work, what these changes consist of has been subject to analysis.

It is evident that the information about semantic change and the new meanings that are present in the database must be compatible with the description model that is already established in SO. According to Svensén (2009: 211–212), the polysemy structure of the words in a dictionary can be described linearly, i.e., as a number of discrete units arranged in a sequence. This observation is primarily valid for monolingual Swedish dictionaries like *Natur och Kulturs stora svenska ordbok* (Köhler & Messius 2006) and *Bonniers svenska ordbok* (Sjögren & Györki 2010). The same is valid for many dictionaries of English, like the *Longman dictionary of contemporary English* and the *Merriam-Webster dictionary*. However, the polysemy structure of a word can also be described in terms of a limited number of main/core senses, to which groups of subsenses/shades are associated (i.e., in an hierarchical order of senses). This is the approach that has been adopted in SO, where the way in which the subsense(s) are related to the main sense is also explicitly specified. This principle is also valid in *Den Danske Ordbog* ("DDO"), the most comprehensive monolingual dictionary of contemporary Danish. To illustrate how these two different principles work in these dictionaries, consider the example *ansiktslyftning* 'face-lift'.[1] First of all, the word refers to surgery, but it is also used metaphorically to refer to repairs that make a building, for example, look newer or better. When the polysemy structure is described linearly in the dictionary, these two meanings are listed as 'meaning 1' and 'meaning 2'. However, when the polysemy structure of the dictionary is hierarchical, the two meanings are listed as 'meaning 1' and 'meaning 1a', because the second meaning is considered to be a metaphorical subsense of the main sense ('meaning 1').

In the hierarchical variant, the relationship between the senses is typically categorized in terms of meaning extension, meaning specialization, metaphorical (figurative) use, etc. An example of meaning extension can be seen in the meaning of the noun *visitkort* 'visiting card'/'business card'. *Visitkort* originally only referred to a kind of concrete paper card with printed information (name, company name, address, etc.). Nowadays, we can find digital visiting cards as well. The process of meaning specialization can be exemplified by the noun *nedtrappning* cf. 'tapering', referring to a gradual decrease of something. The noun

---

[1] All English translations in this chapter are by the authors.

has a special meaning in a medical context: it refers to the process of gradually lessening or reducing the use of a medicine, etc. Finally, an example of a semantic change which is based on metaphorical (figurative) use of a existing word is the already-mentioned *ansiktslyftning* 'face-lift'. (See e.g. Malmgren 1988: 181, for a discussion of meaning relations in the lexical database, and the classic Waldron 1967 for important background information in this area.)

Svensén (2009) points out that an important point in arranging the meanings is to determine how far a subsense of a meaning should be allowed to depart from the main sense before the lexicographer has to consider establishing a new and independent sense. Svensén (2009: 212, 363) also concludes that once the division into senses has been implemented, the order in which the senses are to be presented must be determined. Traditionally, dictionaries have applied a historical order, starting with the oldest sense and ending with the most recent (see, e.g., *Svenska Akademiens ordbok*, 'The Swedish Academy dictionary', henceforth "SAOB"). However, an arrangement of this kind has disadvantages, for example, this approach is not suitable for the majority of users, since they might give up before they find the meaning they were looking for.

In summary, a central goal for the editorial team of SO is both to provide a correct description of contemporary Swedish and to show the relationships between different senses. At the same time, the team aims at compiling a lexicographic resource that is an understandable and useful tool for its intended users in different user situations (including the reception and production of Swedish).

3 Examples from the lexicographer's shop floor

In this section, we examine a number of Swedish words that represent well-known types of semantic change, which have been explored in previous research on developments in the Swedish vocabulary during the last decades: changes in concepts and their reference (discussed in Svensén 2009), emotive change (related to, for example, "feminist language change"; see Wojahn 2015: 35–52), changes in constructional behaviour (Malmgren 2003), and, finally, grammaticalisation and pragmaticalisation (see, e.g., Rosenkvist & Skärlund 2011). Our main contribution is to highlight and discuss several lexicographic problems associated with these types of change.

3.1 Concepts, distinctive features, and prototypical meanings

Definitions in dictionaries are associated with word forms or lemmas. According to a common assumption, definitions are assumed to correspond to concepts in

the minds of language users. As normally understood in the field of lexicography, an intensional definition of a concept, the standard format for definitions in general-language dictionaries, consists of a superordinate concept (genus proximum) of the concept to be defined and at least one distinctive feature (differentia specifica) specific to the concept in question (Svensén 2009: 218–221). The distinctive features specify in which respects the concept to be defined differs from other concepts that are related in the same way (subordinated) to the genus proximum. For example, consider the concepts 'quadrilateral', 'rhombus', and 'rectangle'. An intensional definition of 'rhombus' states that genus proximum is 'quadrilateral' and adds one or several features that distinguishes it from the concept of rectangle, e.g., that all sides have equal length. Importantly, the number of features has to be adjusted so that the definition does not become too narrow or too broad (see also Atkins & Rundell 2008: 414–417).

In SO, nouns and verbs are defined in this way. For example, the main sense of *örn* 'eagle' is *typ av stor rovfågel med långa breda vingar, kraftig näbb och grova klor...* ('type of large bird of prey with long broad wings, strong beak and robust claws'). Another example is *bryta*, literally 'break' (c.f. *broken*), where one of the main senses is *tala med främmande uttalsmönster...* ('speak with a foreign accent').

However, the genus-and-differentia model is sometimes unworkable, since large areas of the lexicon do not fit this taxonomic model. Furthermore, the goal of identifying the necessary and sufficient distinctive features of a lexical unit is questionable (see, e.g., Atkins & Rundell 2008: 416). According to prototype theory, it is impossible to determine which distinctive features are both necessary and sufficient in defining a certain category, since the borderlines between categories are fuzzy (Svensén 2009: 224, see also Rosch & Mervis 1975). Consequently, the lexicographer may aim for a typification in the meaning description by analyzing many individual instances of words in a corpora, instead of trying to isolate necessary and sufficient conditions. The dictionary user will then see the definition that is normally or typically the intended one (see Atkins & Rundell 2008: 418 with references and Svensén 2009: 222–223).

But what happens when, for example, a category of concrete objects, which a lexical item refers to, radically changes over time? How does such a change affect the definition of the word in the dictionary? And what happens when a lexical item referring to a certain kind of state or condition appears in new contexts?

The first case can be illustrated by the noun *bil* 'car'. According to SAOB, the word was established more than a hundred years ago and is currently in use today. Whilst there has been a remarkable technological development of cars, has the meaning of the word changed? In the following, we focus on the lexicographical

consequences of the technical development that has taken place with respect to these vehicles. In SO the main sense of the word is described in the following way: *motordrivet, (vanligen) fyrhjuligt fordon med plats för ett litet antal personer och vanligen främst avsett för persontransport* ('motorized, usually four-wheeled vehicle with room for a small number of people and usually primarily intended for passenger transport'). Although the type of fuel has varied over time and that there exist self-driving cars nowadays, the SO definition, with its superordinate concept and distinctive features, is so general that it points out both older and younger car models. In other words, it still points out a typical car and hence the definition does not need to be updated. However, by using language examples like *familjebil* 'family car', *småbil* 'small car', *elbil* 'electric car', *hybridbil* 'hydrid car', *en bensinsnål bil* 'a petrol-efficient car', *en fyrhjulsdriven bil* 'a four-wheel drive car' and *en förarlös bil* 'a self-driving car', the lexicographer indicates that there is a certain range within the concept and that the referents of the noun can be quite different.

The second case, when a lexical item referring to a certain kind of state or condition appears in new contexts, is illustrated by the abstract noun *nollvision* 'vision zero'. When this word was introduced in the 1990s, it was the name of a government safety project, which had the aim that no one should be killed or seriously injured as a result of a traffic accident in Sweden. The name has become a common noun in Swedish, and in SO (2009) it is defined in the following way: *vision som går ut på att ingen ska dödas eller skadas allvarligt i trafiken* ('vision that no one should be killed or seriously injured in traffic'). However, nowadays the noun also appears in other contexts (cf. the process of generalisation in the well known Waldron 1967). The word may also concern societal aims with regards to the number of suicides and cases of domestic violence. In this particular case, the lexicographers of the second edition have to decide whether the definition from 2009 should be (1) reformulated or (2) complemented. The definition from the first edition can, of course, be revised so that it becomes more general. Since SO is a synchronic dictionary, the sense development of the word during the last decades does not have to be shown. Alternatively, the definition from 2009 can be regarded as the main sense, and a subsense referring to the same kind of safety policy in other contexts could be added. According to the usage in modern corpora, the traffic context is the most recurrent and typical. For this reason, it is likely that the editorial team of the second edition of SO will choose the second option.

3.2 The development of a semantic field: Computers and information technology

There is a clear connection between society and vocabulary, as the seminal works Ullmann (1962) and Waldron (1967) point out. The vocabulary is the most flexible part of a language and new words are introduced hand in hand with new things, new ideas. At the same time, words disappear as the social reality that the word refers to changes. *Nyordsboken* (Moberg 2000: 11–12), which elaborates on the ideas of Ullman and Waldron, presents the following five different ways in which the Swedish vocabulary expands:

1. Already existing words or expressions are given a new meaning or a new area of use.

2. Existing words are combined into a new compound word or phrase.

3. A derivation suffix is added to an already existing word.

4. Existing words or expressions are abbreviated.

5. Words or expressions are borrowed from other languages.

Points 1 and 5 are the most relevant in the present context. Frequently, new words are created through a combination of these ways. Most words that have acquired a new meaning in addition to an already existing meaning are semantic loans (borrowed meanings). In such cases, a word that already exists in Swedish acquires a new meaning through the influence of a foreign word.

Several examples of semantic loans can be found in the field of computers and information technology. Swedish words like *virus* 'virus', *mus* 'mouse', *ikon* 'icon', *mapp* 'folder', *portal* 'portal', and *surfa* 'surf' were already established when their new computer-related meanings were borrowed into Swedish. Hence, these words became (if not already) polysemous.

A quick look in the first edition of SO shows that the SO (2009) editorial team have chosen to treat these new meanings in slightly different ways. Consequently, one can discern certain inconsistencies in how the different words were treated. Despite the fact that all the new meanings are metaphorical (figurative), in some cases the new meanings have formed the basis for a special main sense (as in the case of *mus* and *portal*) and sometimes for a sub-sense to the senses already established (as in the case of *virus* and *surfa*). When defining a new sense, the lexicographers have regarded the new sense as semantically remote and separated from the meanings already described (cf. Svensén 2009 in Section 2).

Some verbs with computer-related meanings that are not registered in SO (2009), but which will most likely be added to the forthcoming second edition, are *importera* 'import', *exportera* 'export', and *strömma* 'stream'. In the first two cases, we observe metaphorical but also specialized uses of the words. The new meaning of *strömma* is semantically related to the basic meaning of the same verb, but while the traditional use of the verb is intransitive, the new one can also be used transitively (cf. SOMETHING *strömmar* 'streams' vs. SOMEONE *strömmar* 'streams' SOMETHING). This aspect is also taken into account in the analysis of the verb and may affect how the new meaning is treated in the dictionary.

A slightly different kind of word that has received a marked increase in use, not least through social media, is the noun *hatare* 'hater'. According to SAOB, the word has been used since 1541 in Swedish texts. Without doubt, the traditional meaning and the new use have many semantic features in common, but the context in which the new meaning appears should of course be included in a description of how the word is normally or typically used today.

Recently, new uses of the Swedish verbs *posta* and *texta* (in the senses 'to publish on the internet' and 'to send an sms') have been noticed (cf. the traditional meanings 'to post a letter' and 'to write in block letters'). There is no doubt that these new uses of the words occur in young people's spoken language. The question is, however, whether these uses are sufficiently established. This can be determined by searches in different corpora.

Whether the above observations reflect a change in the meaning of existing words can also be discussed. The semantic difference between the more established and the new uses of *texta* is so striking that it can be argued that *texta*, in the sense 'send an sms', is simply a new word; it is a homonym to *texta* 'to write in block letters' that has appeared in Swedish.[2] According to Ullmann (1962: 59), homonymy refers to the fact that two synchronically different words have the same surface form and polysemy to the fact that one word has two or more different senses (see Atkins & Rundell 2008: 280–282 for a discussion of polysemy and homonymy in English dictionaries).

Another example from the same subject area is *troll*. The Old Swedish word *troll* 'ugly and supernatural being with a tail; usually perceived as hostile to humans' and the English equivalent 'troll, elf' have been used since the end of the 13th century. Since at least 2009, however, we find an identical word in Swedish with the meaning 'Internet troll'. In terms of surface form, the Swedish nouns *troll*

---

[2]Note that the Swedish *posta* 'publish online, typically on a social media website' is often pronounced in a semi-English fashion (but this is not visible in writing). This also affects the lexicographic classification of the word.

'ugly being with tail' and *troll* 'on the Internet' coincide completely. The words have the same pronunciation and inflection. Furthermore, they have specific semantic aspects in common. For this reason, one could argue that *troll* is polysemous and that the latter use has evolved from the former. The new *troll* is then an example of semantic change. However, the two nouns have completely different origins in that *troll* 'being' is derived from the ancient Swedish *trul, trol*, whereas *troll* 'on the Internet' comes from the English verb and noun *troll* with origins in 'to fish by trolling' (https://www.merriam-webster.com/dictionary/trolling). For this reason, it is more reasonable to consider *troll/troll* as homonyms instead of arguing that the word is polysemous. However, as the widely-consulted Lyons (1968: 406) states:

> The distinction between homonymy and multiple meaning is, in the last resort, indeterminate and arbitrary. Ultimately, it rests upon either the lexicographer's judgement about the plausibility of the assumed 'extension' of meaning and some historical evidence that the particular 'extension' has in fact taken place.

Lyons (1968: 406) also points out that "the arbitrariness of the distinction between homonymy and multiple meaning is reflected in the discrepancies in classification between different dictionaries" (cf. Svensén 2009 in Section 2).

3.3 Emotive meaning

The cases above concern cognitive (or denotative) meaning. There are, however, also dimensions of emotive (or connotative) meaning in language (see Svensén 2009: 214 for a discussion of different notions of meaning). A natural starting point for lexicographers is Stevenson's famous discussion of this notion:

> The emotive meaning of a word is a tendency of a word, arising through the history of its usage, to produce (result from) affective responses in people. It is the immediate aura of feeling which hovers about a word. Such tendencies to produce affective responses cling to words very tenaciously. It would be difficult, for instance, to express merriment by using the interjection "alas". Because of the persistence of such affective tendencies (among other reasons) it becomes feasible to classify them as "meanings". (Stevenson 1937: 23).

Stevenson claims that the emotive meaning of a given word arises "through the history of its usage". The emotive aspects of meaning are closely intertwined

with attitudes towards the concept referred to; it is therefore expected that the emotive meaning can change over time. Consider the word *democracy*, discussed by Stevenson. Today, the word has a positive emotive meaning, but one can imagine that democratic forms of government might fall out of popularity: the word would then keep its cognitive meaning but change its emotive (Stevenson 1944: 72).

In present-day Swedish, consider the examples *lapp* 'Lapp', *eskimå* 'Eskimo', and *indian* 'American Indian'. These words were previously stylistically neutral in Swedish. Today, however, they are not appropriate in newspaper texts and similar genres. The perception of these words has clearly changed in recent history. In all these three cases, the viewpoint of the ethnic group is today a relevant factor for most speakers and writers of public discourse. Notwithstanding this concern, it might not be obvious what the view of the relevant group is. In the case of *lapp*, the denoted ethnic group (the Sami of northern Scandinavia) perceive the expression as strongly derogatory,[3] but the viewpoints of the groups referred to by the Swedish expressions *eskimå* and *indian* are not as obvious. In the latter two cases there exist disagreements about proper labels within the groups (Bird 1999; see the entry on *eskimåer* in the *Nationalencyklopedin*[4]). Furthermore, such disagreements are often embedded in larger complex cultural and social debates in North America and South America (in the case of the English *Indian*, which is, of course, clearly related to the Swedish word but, nevertheless, a different one), and in Denmark and Greenland (in the case of Swedish *eskimå*, closely related to the Danish *eskimo*). These debates and discourses influence the public discourse in Sweden, and the Swedish language, but the distance from the debates increases both the variation in emotive charge across the population of Swedish speakers and the felt complexity, or perhaps unclarity, in relation to the emotive meaning of these words. In the forthcoming second edition of SO, *lapp* has the usage marker *starkt nedsättande* ('strongly derogatory'). *Eskimå* and *indian* have the usage marker *kan uppfattas som nedsättande* ('can be perceived as derogatory'), which highlights the variation and complexity of the words' emotive charge.

Other words, not associated with ethnicity, have also changed in emotive meaning. Consider *bög* 'male homosexual' and *flata* 'female homosexual', in the semantic field of sexual orientation. These words used to be clearly derogatory; now they have been partly reclaimed by the LGBTQI-community and can be used with a neutral emotive meaning. The meanings are, however, context dependent in these cases, and the conditions for application are quite complex: the emotive

---

[3]See e.g. Sametinget (2017: 5).

[4]https://www.ne.se/uppslagsverk/encyklopedi/l%C3%A5ng/eskim%C3%A5er

charge depends on, for instance, tone of voice, discourse topic, and the identity of the speaker (cf. Petersson & Sköldberg 2020).

It is difficult to determine when a word has changed its emotive meaning. The issue is pressing, since such modifications can occur rapidly and words quickly can become controversial in the public sphere. Debates about a word in newspapers and on social media can be indicators, but the intuitions of the lexicographer play a role as well. In Section 4, we discuss whether (and how) automatic methods can be of help in detecting changes in emotive meanings.

3.4 Constructional behaviour

Another topic in the field of semantic change concerns different kinds of constructions and word combinations. It is well-known that a language user's mastery of the syntagmatic properties of lexical items (including relevant collocations) has major consequences for whether their language is perceived as idiomatic or not. Research has also shown that even advanced learners have difficulties with the use of different kinds of conventionalized expressions in their second language (see, e.g., Nation 2013: 479 with references). SO aims to account for this theme, in addition to providing a complete description of headwords and their syntagmatic properties.

A word that has evolved considerably in recent decades is the reflexive verb *gifta sig* 'get married'. The main sense of the word, 'enter into marriage', has been established since the 14th century and, for centuries, the constructions for the word were SOMEONE *gifter sig* 'gets married' (with SOMEONE) or SOME PEOPLE *gifter sig*. Changes in society have had consequences for this verb, however. Since same-sex marriages regularly take place in society, the gender of the referents of SOMEONE and SOME PEOPLE, the subject and the object, has undergone change. However, because the definition of the verb is *ingå äktenskap* 'to marry' which includes both classes of marriage, it has not been revised in the (forthcoming) second edition of the dictionary. The relatively new situation, where the referents of the subject and object can be of the same gender can be illustrated by the following language example in a dictionary article: *hon har gift sig med sin fästmö* 'she has married her fiancée[F]'.

Since the 1990s, there is also a metaphorical use of the same verb. In corpora, we can find examples such as *låt såsen dra i några timmar innan servering så att smakerna hinner gifta sig* 'let the sauce soak/draw for a few hours before serving so that the flavors have time to get married' with the constructions SOMETHING *gifter sig* (with SOMETHING) and SOME THINGS *gifter sig*. This sense was well-established when the first edition of SO was compiled, but, due to a lack of more

advanced tools, it was not noticed by the lexicographers. By using corpus tools (for instance the ones provided by the research unit Språkbanken Text, see *Korp* (Borin et al. 2012)) it is now much easier for lexicographers to register this type of semantic variation.

The current search interface of Korp (Borin et al. 2012) includes three independent ways of viewing the results of a search. These are (i) the KWIC concordance view, (ii) the statistics view, and (iii) the Word Picture view.

According to Atkins & Rundell (2008), a KWIC concordance is a basic corpus lexicography tool. Right-sorted and left-sorted concordances often give a "powerful, visual representation of a word's recurrent patterns – in a way that is impossible to ignore or overlook" (Atkins & Rundell 2008: 105). By using the statistics view, it is, for example, relatively simple to compare the frequency of different spelling variants of a word.

Finally, to further develop the treatment of collocations in a dictionary, the Word Picture view is very useful. The Word Picture view in Språkbanken, which is based on the association measure lexicographer's mutual information, gives an overview of selected syntactical environments of a word (i.e., typical verbs, prepositions, pre-modifiers, and post-modifiers) (see Borin et al. 2012: 476; cf. the word sketches generated by Sketch Engine, a corpus tool presented in Kilgarriff et al. 2004). Consequently, by using Word Picture, the lexicographer can provide a more comprehensive description of the semantics of headwords (like *gifta sig*) and their phraseological behaviour.

Another kind of example concerns collocations including the word *gripa* 'profoundly touch, move, affect'. In his 2003 diachronic study, Malmgren states that the verb has become more frequent in abstract transitional phrases, such as *gripas av förtjusning/misströstan/raseri/svårmod* 'be affected by delight/despair/rage/discouragement'. At the same time, the use of the verb *falla* 'fall, yield, give way to' in corresponding phrases has become less frequent during the 20th century. Older uses such as *falla i frestelse/förtjusning/misströstan/raseri* 'give way to temptation/delight/despair/rage' are now perceived as obsolete (Malmgren 2003: 140–141). We thus observe that some collocation verbs have become obsolete and are replaced by other verbs. In other words, some verbs (like *gripa*) demonstrate expanded or extended combinatorial properties while other verbs (like *falla*) are subject to reduced and more limited combinatorial properties. Tools like Word Picture can be useful for future studies in this area.

3.5 Pragmaticalisation

Beeching (2010) explores pragmaticalisation, a process closely related to grammaticalisation (Traugott & Dasher 2001). Pragmaticalisation takes place when a

content expression develops into a pragmatic marker, in contrast to grammaticalisation, which concerns the development into purely grammatical functions. Beeching focuses on the English *effectively* and the French *effectivement* and shows that the expressions have developed from the shared meanings 'efficaciously' and 'in fact' to different pragmatic meanings in the two languages: in French 'that is so' (used as an answer to a question), in English 'contrary to experience' or a purely hedging meaning (expressing uncertainty about the speaker's assertion). The explanation is related to Traugott & Dasher (2001), where a theory of grammaticalisation in terms of conversational implicatures is put forward (see the classic in pragmatics Grice 1975); in short, repeated implicatures can over time become integrated parts of semantic meanings. In the English case, the meaning 'in fact' and 'contrary to experience' invite the inference that the speaker is not making a certain assertion.

Related examples in Swedish are *typ* 'type' and *exakt* 'exactly'. These expressions have developed from content words with clear cognitive (denotative) meanings to pragmatic markers (discourse particles). Rosenkvist & Skärlund (2011) shows how *typ* develops from a noun (*en envis typ* 'a stubborn kind of fellow') via a two-word preposition *av typ* (*en båt av typ lyxjakt* 'a boat of a luxury type'), to a preposition (*Han gillar musik typ Dylan* 'He likes music of Dylan's kind') and then, in recent history, to a pragmatic marker used for hedging (*Välkommen till England, typ* 'Welcome to England, or whatever'). *Exakt* is used as an interjection, affirming previous statements (parallel to the pragmatic function of the English *exactly*).

In the dictionary, all of these uses should be described. However, SO is faced with a number of challenges, with regards to the examples reported on here. First, the different uses of *typ* and *exakt* are related, but the standard labels for meaning relations, which concern mechanisms of metaphor, generalisation, and similar types of change, are not suitable in this context (see Section 2). A new set of labels for pragmatic meaning relations is called for. Second, the structure of the headwords and main senses in SO treats the different uses of *typ* and *exakt* as different headwords, since they differ in word class. It is debatable whether a strict adherence to principles concerning the structure of the dictionary is relevant here; perhaps a more user-friendly approach would be to list all uses of *typ*, and all uses of *exakt*, under the same headword (see Section 2).

4 Discussion

In this chapter, we have discussed semantic changes in Swedish words from a lexicographic perspective. The starting point for the reasoning has been, first

and foremost, the work conducted by the editorial team for the forthcoming second edition of SO, a comprehensive synchronic dictionary with emphasis on the semantics of the lexical units.

In the chapter, we have discussed a number of different kinds of semantic changes. The change can consist of a certain word taking new meaning (e.g. *nollvision* 'vision zero'). Sometimes, one is able to identify a similar development in a group of words belonging to the same semantic field (e.g., *virus* 'virus', *strömma* 'stream', etc.). The semantic change can also consist of a word being associated with more negative emotive meaning in public discourse (e.g., *indian* 'American Indian'). Such changes might happen fairly quickly. Furthermore, semantic change can consist of a changed constructional behaviour of a word. The referents of the subject and the object of a verb might shift (as in *gifta sig* 'get married') and the tendency of a word to be included in collocations may increase or be reduced (*gripa* 'profoundly touch, move, affect', and *falla* 'fall, yield, give way to'). A word may also, by pragmaticalisation, lose its lexical meaning and become a function word (e.g., *typ* 'kind of'). It is also the case that a semantic change of a word may be relatively established among some language users and within a certain kind of language (e.g., spoken youth language) but it may be relatively unknown among other groups of language users. This is the case for *posta* 'post (on blogs etc.)'. Finally, it may be the case that what one initially might have thought was a new meaning of an established word is, in fact, a new word, i.e., a homonym (as demonstrated in our discussion of *troll* 'Internet troll'). In summary, the phenomenon of semantic change, in Swedish and other languages, is multifaceted and diverse. But regardless of the type of change one examines, all types are relevant to lexicographers because these changes should lead to revisions of dictionary articles.

The development of more formal, computational tools for discovering semantic changes is most welcome. In closing, we share some thoughts on how such methods can assist us in achieving our aims with SO.

First, given the fact that SO aims to reflect general vocabulary, we are primarily interested in changes in the general vocabulary of Swedish and not in developments of meaning in technical language. Notwithstanding this stated aim, we would like to obtain more information about the differences in the semantics and the usage of Swedish words in newspaper language and in social media, for example. Our major area of interest determines which materials we should examine. However, we fear that the Swedish corpora available at the moment are too limited, and we thus propose that the existing Swedish corpora, especially with regards to the inclusion of newer texts, need to be radically improved.

Second, lexicographers have traditionally focused on written language. The main focus of lexicographers has been, and remains, on the description of established (lexicalized) changes with a relatively good spread in different corpora. However, we would also like to have more data on spoken language. Although perhaps expensive and practically difficult, a point on our wish list is a searchable corpus of authentic spoken dialogues.

Third, several semantic changes can be identified by use of Word Picture in Korp and similar technologies. It is clear that such technology is significantly helpful for observing metaphorical and metonymical changes, and for specifications and generalisations as well. However, it should be noted that any analysis of the data that is provided by different corpus tools is highly dependent on the lexicographer's linguistic intuitions and experience in the field.

Fourth, it seems to be the case that certain semantic changes cannot be identified by technologies such as Word Picture or other automatically generated information about linguistic contexts. Emotive meanings are especially difficult to identify using such techniques. In these cases, debates about words, in public discourse and social media, play a pivotal role, but the lexicographer's linguistic intuitions are crucial as well. Language technology can provide useful information about the genres, where controversial words are discussed and written about. For instance, if words related to minority groups are increasingly used on social media, that would be useful for us to know.

Fifth, similar difficulties arise in cases of pragmaticalisation and the related process of grammaticalisation. It is unclear to us how Word Picture or a similar tool would be of use in these regards. However, we could start from the problems and questions of lexicography and list a number of items that we would like to keep track of with automatic methods. This list would then include cases like *typ* and *exakt*, where it is clear that pragmaticalisation has taken place.

Finally, we claim that computational methods may be of use in studies of collocations. By examining changes over time in the narrower context of words, one can register new meanings and uses of verbs like *gifta sig* 'get married', *gripa* 'profoundly touch, move, affect' and *falla* 'fall, yield, give way to'. See, e.g., the study of variations in bigrams over time in Nimb et al. (2020), where a method of updating headwords in DDO with new semantic information is investigated. Their study, which combines corpus statistics with manual annotations, is based on "the hypothesis that the variation in bigrams over time in a corpus might indicate changes in the meaning of one of the words" (Nimb et al. 2020: 112). Furthermore, the fact that verbs such as *exportera, importera*, and *strömma* and the noun *hatare* are now used in computer contexts should be discernible if one compares broader contexts in corpora reflecting language from different periods of

time. This can be related to Cook et al. (2013), who, based on an automatic word sense induction system, compare three sentence contexts of target words in two corpora representing different language periods, and evaluate whether there are any differences in usage of the target words.

We also suggest that language technology may be relevant to contrastive studies. Researchers studying different languages can benefit from each other's work. For example, in the field of computers and information technology, one can see clear parallels between the development of different words in English and in Swedish (see Section 3.2). Slightly simplified, if, for example, the English word *virus* begins to be used metaphorically about computers in English, it is not surprising if the same development in the corresponding Swedish noun is observed. In this context, computational methods would be most welcome.

From our perspective of lexicography, the point of computational methods is to provide sharper tools and allow for a more precise and formal methodology. In practice, language technologists might provide lexicographers with candidate lists of lexical items that seem to have undergone a semantic change. These data sets could then be assessed by lexicographers (see the methods in e.g. Cook et al. 2013 and Nimb et al. 2020). The methods for detecting semantic change would then, hopefully, become more precise. The production of dictionaries would also become more systematic and less reliant on the subjective judgment of individual lexicographers (see Cook et al. 2013: 50).

A pertinent issue in this discussion is deciding on which semantic changes should be prioritized. All of the cases discussed in this chapter are of interest, from a linguistic point of view. But for the dictionary user, especially second-language learners, the most important examples are, perhaps, the examples with negative, or unclear, emotive (connotative) meanings. Therefore, automatic methods that could help us improve the lexicographic quality with regards to the emotive aspects of words, and the changes in emotive meaning, would be most welcome.

Abbreviations

SO   The contemporary dictionary of the Swedish Academy
GU   The University of Gothenburg
SAOB   The Swedish Academy dictionary
KWIC   Keyword-in-context

References

Atkins, Sue & Michael Rundell. 2008. *The Oxford guide to practical lexicography*. Oxford: Oxford University Press.

Beeching, Kate. 2010. Semantic change: Evidence from false friends. *Languages in Contrast* 10(2). 139–165.

Bird, Michael Yellow. 1999. What we want to be called: Indigenous people's perspectives on racial and ethnic identity labels. *American Indian Quarterly* 23(2). 1–21.

Borin, Lars, Markus Forsberg & Johan Roxendal. 2012. Korp: The corpus infrastructure of Språkbanken. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Mehmet Uğur Doğan, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of LREC 2012*, 474–478. Istanbul: ELRA. https://spraakbanken.gu.se/korp.

Cavallin, Karin. 2012. Exploring semantic change with lexical sets. In *Proceedings of the XV EURALEX international congress*, 1018–1022. Oslo: EURALEX.

Cook, Paul, Jey Han Lau, M. Rundell, Diana McCarthy & Timothy Baldwin. 2013. A lexicographic appraisal of an automatic approach for detecting new word-senses. In I. Kosem, J. Kallas, P. Gantar, S. Krek, M. Langemets & M. Tuulik (eds.), *Proceeedings of eLex 2013*, 49–65. Ljubljana/Tallinn: Trojina, Institute for Applied Slovene Studies/Eesti Keele Instituut.

Cook, Paul, Jey Han Lau, Michael Rundell, Diana McCarthy & Timothy Baldwin. 2014. Applying a word-sense induction system to the automatic extraction of diverse dictionary examples. In *Proceedings of the 16th EURALEX International Congress*, 319–328. Ljubljana: Ljubljana University Press.

Det Danske Sprog- og Litteraturselskab (ed.). 1918–. *Den Danske Ordbog*. https://ordnet.dk/ddo.

Grice, Herbert Paul. 1975. Logic and conversation. In P. Cole & J. Morgan (eds.), *Syntax and semantics*, vol. 3, Speech acts, 41–58. New York: Academic Press.

Kilgarriff, Adam, Husák Miloš, Katy McAdam, Michael Rundell & Rychlý Pavel. 2008. GDEX: Automatically finding good dictionary examples in a corpus. In E. Bernal & J. DeCesaris (eds.), *Proceedings of the XIII EURALEX international congress*, 425–433. Barcelona: Universitat Pompeu Fabra.

Kilgarriff, Adam, Pavel Rychlý, Pavel Smrz & David Tugwell. 2004. The sketch engine. In *Proceedings of the 11th EURALEX International Congress*, 105–115. Lorient: Université de Bretagne Sud.

Köhler, Per Olof & Ulla Messius. 2006. *Natur och Kulturs stora svenska ordbok*. Christian Mattsson (ed.). Stockholm: Natur och Kultur.

Lyons, John. 1968. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press.

Malmgren, Sven-Göran. 1988. On regular polysemy in Swedish. In *Studies in computer-aided lexicology* (Data Linguistica 18), 179–200. Stockholm: Almqvist & Wiksell.

Malmgren, Sven-Göran. 2003. Begå eller ta självmord? Om svenska kollokationer och deras förändringsbenägenhet [Commit or take suicide: On Swedish collocations and their tendency to change]. *Språk & Stil* 13. 123–168.

Merriam-Webster (ed.). 2019. *Merriam-Webster dictionary*. Version 5.2.0. Springfield: Merriam-Webster.

Moberg, Lena. 2000. *Nyordsboken*. Stockholm: Svenska språknämnden.

Nation, Paul. 2013. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nimb, Sanni, Nicolai Hartvig Sørensen & Henrik Lorentzen. 2020. Updating the dictionary: Semantic change identification based on change in bigrams over time. *Slovenščina 2.0* 8(2). 112–138.

Pearson Longman (ed.). 2009. *Longman dictionary of contemporary English*. 5th. Harlow, UK.

Petersson, Stellan & Emma Sköldberg. 2020. To discriminate between discrimination and inclusion: A lexicographer's dilemma. In *EURALEX XIXI book of proceedings Vol. 1*. 381–386. Alexandroupolis: EURAC.

Pilán, Ildikó. 2016. Detecting context dependence in exercise item candidates selected from corpora. In *Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications*, 151–161. San Diego: ACL.

Rosch, Eleanor & Carolyn B. Mervis. 1975. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7. 573–605.

Rosenkvist, Henrik & Sanna Skärlund. 2011. Grammatikalisering i nutid – utvecklingen av *typ* fram till 2009 [Present-day grammaticalization – the development of *typ* until 2009]. *Språk & Stil* 21. 5–25.

Sametinget. 2017. *Svenska Sametingets kommentar till Sveriges 22:a och 23:a periodiska rapport till Kommittén för avskaffande av rasdiskriminering* Comments by the Swedish Sami Parliament regarding Sweden's 22nd and 23rd periodic reports to the Committee on the Elimination of Racial Discrimination. Kiruna/-Giron: Sametinget. https://www.sametinget.se/117555.

Sjögren, Peter A. & Iréne Györki. 2010. *Bonniers svenska ordbok*. Stockholm: Bonnier.

Stevenson, Charles Leslie. 1937. The emotive meaning of ethical terms. *Mind* 181. 14–31.

Stevenson, Charles Leslie. 1944. *Ethics and language.* New Haven: Yale University Press.

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

Svenska Akademien. 2009. *Svensk ordbok utgiven av Svenska Akademien.* https://svenska.se.

Svenska Akademien. 1898–. *Svenska Akademiens ordbok.* https://svenska.se/.

Traugott, Elizabeth Closs & Richard B. Dasher. 2001. *Regularity in semantic change.* Cambridge: Cambridge University Press.

Ullmann, Stephen. 1962. *Semantics: An introduction to the science of meaning.* Oxford: Basil Blackwell.

Waldron, Ronald A. 1967. *Sense and sense development.* London: Andre Deutch Limited.

Wojahn, Daniel. 2015. *Språkaktivism: Diskussioner om feministiska språkförändringar i Sverige från 1960-talet till 2015* [Language activism. Discussions on feminist language change in Sweden from the 1960s until 2015] (Studier utgivna av Institutionen för nordiska språk vid Uppsala universitet 92). Uppsala: Institutionen för nordiska språk, Uppsala universitet.