Using language technology resources and tools to construct Swedish FrameNet

Dana Dannélls   **Karin Friberg Heppin**   **Anna Ehrlemark**

Department of Swedish
University of Gothenburg
`firstname.lastname@svenska.gu.se`

Abstract

Having access to large lexical and grammatical resources when creating a new language resource is essential for its enhancement and enrichment. This paper describes the interplay and interactive utilization of different language technology tools and resources, in particular the Swedish lexicon SALDO and Swedish Constructicon, in the creation of Swedish FrameNet. We show how integrating resources in a larger infrastructure is much more than the sum of the parts.

1 Introduction

This paper describes how Swedish language technology resources are exploited to construct Swedish FrameNet (SweFN),[1] a lexical-semantic resource that has been expanded from and constructed in line with Berkeley FrameNet (BFN). The resource has been developed within the framework of the theory of Frame Semantics (Fillmore, 1985). According to this theory, semantic frames including their participants represent cognitive scenarios as schematic representations of events, objects, situations, or states of affairs. The participants are called frame elements (FEs) and are described in terms of semantic roles such as AGENT, LOCATION, or MANNER. Frames are evoked by lexical units (LUs) which are pairings of lemmas and meanings.

To get a visualization of the notion of semantic frames consider the frame `Vehicle_landing`. It has the following definition in BFN: "A flying VEHICLE comes to the ground at a GOAL in a controlled fashion, typically (but not necessarily) operated by an operator." VEHICLE and GOAL are the *core elements* that together with the description uniquely characterize the frame. Their semantic types are *Physical_object* and *Location*. The *non-core elements* of the frame are: CIRCUMSTANCES, COTHEME, DEGREE, DEPICTIVE, EVENT_DESCRIPTION, FREQUENCY, GOAL_CONDITIONS, MANNER, MEANS, MODE_OF_TRANSPORTATION, PATH, PERIOD_OF_ITERATIONS, PLACE, PURPOSE, RE_ENCODING, SOURCE, and TIME. The lexical units evoking the frame are: *land.v, set_down.v*, and *touch_down.v*. In addition, the frame contains a number of example sentences which are annotated in terms of LUs and FEs. These sentences carry valence information about different syntactic realizations of the FEs and about their semantic characteristics.

Currently SweFN contains around 1,150 frames with over 29,000 lexical units of which 5,000 are verbs, and also 8,300 semantically and syntactically annotated sentences, selected from a corpus.

SweFN has mainly been created manually, but as a response to an ever increasing complexity, volume, and specialization of textual evidence, the creation of SweFN is enhanced with automated Natural Language Processing (NLP) techniques. In contrast to the construction of English resources, as well as the construction of framenets for other languages, the resources used to construct SweFN are all linked in a unique infrastructure of language resources.

2 The development of framenets in other languages

FrameNet-like resources have been developed in several languages and have been exploited in a range of NLP applications such as semantic parsing (Das et al., 2014), information extraction (Moschitti et

[1]`http://spraakbanken.gu.se/eng/resource/swefn`

al., 2003), natural language generation (Roth and Frank, 2009), and semi-automatic disambiguation of polysemous words (Alonso et al., 2013).

Currently the most active framenet research teams are working on Swedish FrameNet (SweFN) (Borin et al., 2010; Heppin and Gronostaj, 2014), Japanese FrameNet (JFN) covering 565 frames, 8,500 LUs, and 60,000 annotated example sentences (Ohara, 2013) and FrameNet Brazil (Br-FN) for Brazilian Portuguese (Torrent, 2013) covering 179 frames, 196 LUs, and 12,100 annotated sentences.[2]

Even though the point of departure for all FrameNet-like resources is BFN, they differ in a number of important aspects. SweFN has focused on transferring frames and populating them with LUs. For each frame there are annotated example sentences extracted from corpora. Sentences illustrate the instantiation of a number of LUs and FEs with regard to the frame, but many LUs do not yet have associated example sentences. BFN and Spanish FrameNet (Subirats, 2009) also use isolated corpus sentences for annotation while the SALSA project for German (Burchardt et al., 2009) has the aim of creating full-text annotation of a German corpus. JFN, Spanish FrameNet, and FN-Br all use BFN software to construct frames, while SweFN uses its own software and tools. Even though JFN uses BFN software and annotations tools for as much compatibility with BFN as possible, the Japanese writing system differs considerably from that of English, and several modifications have been necessary to handle the different character systems and word boundary issues.

Most framenets have the intention of covering general language. However, there are domain specific resources such as, the Copa 2014 FrameNet Brasil, a multilingual resource for the language of soccer and tourism (Torrent et al., 2014) covering Portuguese, English and Spanish. Bertoldi and de Oliveira Chishman (2011) describe work buiding a FrameNet-like ontology for the language of criminal justice contrasting the differences between English and Portuguese languages and legal cultures.

3 Lexical and grammatical resources and tools for Swedish

Swedish FrameNet is part of SweFN++, a larger project with the goal to create a multifaceted panchronic lexical macro-structure for Swedish to be used as an infrastructure component for Swedish language technology and development of NLP applications and annotated corpora. One goal of SweFN++ is to re-use and enhance existing in-house and external lexical resources and harmonize them into a single macro-structure for processing both modern and historic Swedish text (Borin et al., 2010). Another goal is to release all SweFN++ resources under an open content license.

3.1 SALDO – association lexicon

SALDO (Borin et al., 2013a)[3] is a Swedish association lexicon which contains morphological and lexical-semantic information for more than 131,000 entries, of which around 10% are verbs. SALDO entries are arranged in a hierarchical structure capturing semantic closeness between lexemes. Each lexical entry of SALDO has a unique identifier. Each lexical entry, except 41 top nodes, also has a main descriptor, which may be complemented with a second determinative descriptor. These descriptors are other, more central, entries from SALDO. The SALDO entry for the noun *flaska* 'bottle', with its descriptors, is shown in figure 1.

SALDO is the pivot of all the Swedish lexical language technology resources maintained at Språkbanken. Having one pivot resource makes it possible for all Språkbanken resources to be compatible with each other (Borin and Forsberg, 2014).

3.2 Swedish Constructicon

The Swedish Constructicon (SweCcn)[4] is an electronic database of Swedish constructions (Lyngfelt et al., 2012; Sköldberg et al., 2013). Just as it is precursor the Berkeley Constructicon,[5] it builds on experiences from Construction Grammar and is historically, methodologically and theoretically closely related to Frame Semantics and FrameNet (Fillmore et al., 2012). While framenets map single lexical

---

Saldo (1)

Sense	Lemgram	POS	Primary	Secondary	Children (primary)	Children (secondary)
flaska 🐦	*flaska (noun)* (186315)	noun	*förvara*	*hälla*	• *butelj* – • *champagneflaska* champagne • *engångsflaska* engångs Show all (23)	• *back³* förvara • *dryck²* dryck • *flaskborste* borste Show all (19)

Figure 1: A search for the noun *flaska* 'bottle' in SALDO shows that it only has one sense. We are also shown the lemma, the part of speech, the primary descriptor *förvara* 'store.v', the secondary descriptor *hälla* 'pour.v', and finally primary and secondary children, that is entries which have *flaska* as primary or secondary descriptor.

units to the frames they evoke, a constructicon deals with the pairing of form and meaning in more complex linguistic units, typically (partially) schematic multiword units that cannot easily be referred to by either grammatical or lexicographic descriptions alone.

In SweCcn each construction is described individually in a construction entry, defined by its specific characteristics in form, meaning, function, and distribution. Each entry includes a free text definition, schematic structural description, definitions of construction elements (CEs) and annotated example sentences. Since the constructicon must account for both form and meaning, the construction elements can be both semantic roles and syntactic constituents. For example, the construction `reflexiv_resultativ`, instantiated in *äta sig mätt* 'eat oneself full', is defined as a verb phrase where somebody (ACTOR) or something (THEME) performs an action (ACTIVITY) that leads to a result which affects the ACTOR/THEME, expressed with a reflexive particle. The construction roughly means "achieve something by V-ing", and can be applied to both transitive and intransitive verbs, altering the verbs' inherent valence restrictions. The syntactic structure of the construction is [V refl AP], and the construction elements are defined as the semantic roles ACTOR, THEME, ACTIVITY and RESULT, as well as the reflexive particle. Example sentences like *dricka sig full* 'drink oneself drunk' and *springa sig varm* 'run oneself warm' are added to the entry, while an example like *känna sig trött* 'feel tired' does not fit since one doesn't get tired by feeling.

Swedish Constructicon is developed as an extension of Swedish FrameNet and forms a part of the SweFN++ infrastructure. Swedish Constructicon currently consists of about 300 construction entries, ranging from general linguistic patterns to partially fixed expressions, of which a significant part are constructions in the borderland between grammar and lexicon, commonly neglected from both perspectives.

3.3 Karp – open lexical infrastructure

Karp is an open lexical infrastructure with three main functions: (1) support the creation, curation, and mutual integration of the lexical resources of SweFN++; (2) publish all lexical resources at Språkbanken, making them searchable and downloadable in various formats such as Lexical Markup Framework (LMF) (Francopoulo et al., 2006), and Resource Description Framework (RDF) (Lassila and Swick, 1999); (3) offer advanced editing functionalities with support for exploitation of corpora resources (Borin et al., 2013b).

There are 21 resources with over 700,000 lexical entries available in Karp. Since all resources utilize the lexical entries of SALDO, a large amount of information becomes accessible when performing simple searches. For example when we look up the SALDO entry *flaska..1* 'bottle', we find information about the synset from Swesaurus,[6] a WordNet-like Swedish resource, as well as synset and sense from Princeton WordNet,[7] syntactic valence from PAROLE,[8] identifier from Loan Typology Wordlist (LWT),[9]

---

[6]http://spraakbanken.gu.se/eng/resource/swesaurus
[7]http://wordnet.princeton.edu/
[8]http://spraakdata.gu.se/parole/lexikon/swedish.parole.lexikon.html
[9]http://lingweb.eva.mpg.de/cgi-bin/ids/ids.pl?com=simple_browse&lg_id=187

the lexical ID from Lexin,[10] etc. Each of these resources is in turn linked to mono- and multi-lingual information that can be exploited by any other resource or application.

3.4 Korp – Swedish corpora

Korp is a Swedish corpus search interface developed at Språkbanken. It provides access to over 1.6 billion tokens from both modern and historic Swedish texts (Borin et al., 2012; Ahlberg et al., 2013). The interface allows advanced searches and comparisons between different corpora, all automatically annotated with dependency structure using MaltParser (Nivre et al., 2007).

One functionality provided by Korp is *Related Words*. This shows a list of words fetched from SALDO which are semantically related to the search term. Only words that actually occur in the corpora are retrieved by this function. By clicking on one of these, a new corpus search is done with this word as search term (Borin et al., 2012). Another functionality in Korp is *Word Picture* which uses statistical data to select typical examples illustrating collocational semantic relations for chosen expressions. This query system extracts frequent collocations of the word in question along with an analysis of the parts-of-speech of the collocating words.

4 The development of SweFN

As described by the BFN research team, manual construction of a framenet resource involves several steps, including defining frames and frame elements, collecting appropriate lexical units for the frames, comparing the findings with printed dictionaries, extracting syntactic and collocational contexts to illustrate the frame, and analyzing sentences to explore the use of LUs (Fillmore et al., 2003).

The work procedure of SweFN is based on transfer of information from BFN. To a large extent we follow the BFN development process, but the development of SweFN differs in three crucial aspects: (1) when we transfer frames from BFN to Swedish, there is usually no need to re-define them. However, the frames are checked for compatibility with Swedish language and culture; (2) our inventory of LUs is derived from the SALDO lexicon; (3) we utilize in-house resources, all linked in the Swedish infrastructure for language technology, SweFN++.

Taking BFN as a starting point saves time and effort in developing frames. Most of the effort goes to figure out what SALDO entries evoke which frames and to find suitable example sentences. In order to find appropriate LUs evoking a particular frame we consult: (1) the lexical resources in Karp (see section 4.3); (2) printed dictionaries; (3) the corpus infrastructure Korp for concordance search in order to investigate additional uses of the words. This process occasionally results in new frames or modification of the frames of BFN (see section 4.4).

4.1 SALDO

The manual process of constructing a SweFN frame begins with choosing a frame from BFN or word of interest. When we create a frame equivalent to one which already exists in BFN, we transfer the frame features which are more or less language independent from the BFN frame to the SweFN frame. These features include frame description, frame-to-frame relations, and FEs. We then search for appropriate SALDO entries evoking the frame as well as example sentences for annotation. If suitable entries exist in SALDO they are chosen for use as LUs. Otherwise we suggest entries to be added to SALDO (Borin et al., 2013a). Each SALDO sense is allowed to populate only one SweFN frame except in a few cases where some inflectional forms evoke one frame and other forms another frame.

When we instead use a word or expression as a starting point we look up all senses in SALDO and systematically add each sense to the frame it evokes. The selection of LUs from SALDO to populate the frames of SweFN is done in different ways. One method is to determine which of the English LUs of BFN frames have suitable equivalents in Swedish. Thereafter different types of searches are made in SALDO. For example, working on the frame `Containers`, having introduced the noun LU *flaska* 'bottle' one can search for entries ending with *flaska*, thus finding a number of compounds such as *champagneflaska* 'champagne bottle', *droppflaska* 'dropper bottle (med.)', *engångsflaska* (one+time+bottle)

'non-returnable bottle', *glasflaska* 'glass bottle', *halvflaska* (half+bottle) '375ml bottle', *miniatyrflaska* 'miniature bottle', *nappflaska* (pacifier+bottle) 'baby bottle', *sprayflaska* (spray+bottle) 'spray can', *tomflaska* 'empty bottle', *vattenflaska* 'water bottle', *värmeflaska* (heat+bottle) 'warm water bottle', to name a few. Another method is searching for entries having the LU in question as one of the determiners. For example, working on the `Animal` frame, a search may be done on the determiner *djur* 'animal' resulting in a long list of lexical entries for different species of animals, which may be entered into the frame.

The possibility of doing searches in SALDO as described above, in combination with compounding being very productive in Swedish, is one reason for the relatively large number of LUs in SweFN.

4.2 Swedish Constructicon

Constructions are more complex linguistic units than words, they are common in use and difficult to ignore when working with authentic text. One way to enrich SweFN with more representative examples of how to express meaning in language is to include constructions as frame-evoking units in the database. Currently work is being done on systematically linking constructions in SweCcn with frames in SweFN (Ehrlemark, 2014), but the task is not as straight-forward as identifying which frame is evoked by a certain LU. First, not all constructions evoke frames, carrying little meaning from a semantic point of view. This includes such general patterns as constructions for modification, predication, passive voice or filler-gap constructions. Second, constructions that potentially correspond with frames do not always fit the distribution pattern of frame elements described in the target frame. This group includes figurative constructions or constructions that are more, or less, general than the target frame in SweFN. Constructions which do correspond with frames may be called frame-bearing constructions (Fillmore et al., 2012). A frame-bearing construction evokes a target frame in the same manner as an LU, with matching construction elements and frame elements.

The linking of constructions with frames is carried out through manual analysis of constructions and their semantic valence patterns. The work includes paraphrasing the meaning of a construction to identify which frame or frames it may evoke, and thereafter comparing the construction elements with the FEs of the target frame. For example, SweCcn includes three constructions for comparisons: `jämförelse` 'comparison', which has the two subordinate constructions `jämförelse.likhet` 'comparison.similarity' and `jämförelse.olikhet` 'comparison.difference' – all three are Swedish equivalents of corresponding constructions in the Berkeley Constructicon (Bäckström et al., 2014). In all three cases the CEs in the construction entries correspond to the FEs in the `Evaluative_comparison` frame which has the following definition: a PROFILED_ITEM is compared to a STANDARD_ITEM with respect to some ATTRIBUTE. By establishing a link between, in this case the `comparison` constructions and the `Evaluative_comparison` frame, we may enrich the frame with typical example sentences such as *Hennes cykel är bättre än min* 'Her bicycle is better than mine' and *Popband är lika arga som rockband* 'Popbands are as angry as rockbands'.

Another example is the pair of constructions `proportion_i_om` and `proportion_per`, which distinguish different syntactic patterns for expressing proportion in Swedish. In both cases, the construction combines two entities, a numerator and a denominator, joined by a preposition. However, they differ regarding domain of use, preposition used, and definiteness of the second noun phrase. The construction `proportion_i_om` describes time, and therefore corresponds to frames that express proportion in relation to time units, such as `Frequency` and `Speed_description`. The construction `proportion_per` is a more general construction that expresses `Frequency` and `Speed_description` as well as other ratio relations as described in the frames `Relational_quantity`, `Rate_quantification`, `Proportion`, and `Price_per_unit`. Thus, a link between SweFN and SweCcn may refer the user to correct Swedish constructions for ratio relations from the frames they evoke.

At the time of writing, about half of the entries in SweCcn are linked to frames in SweFN. The continuing work with comparing and linking the two resources does not aim to link all constructions with frames, but rather to distinguish frame-bearing from non-frame-bearing constructions. The linking allows the user to easily go between a construction and the frame or frames it evokes and correspondingly

from a frame to constructions evoking the frame. In this way, both SweCcn and SweFN become more representative of the language they set out to describe and better incorporated for future pedagogical and language technological uses.

4.3 Karp

As well as being the editing tool used to build SweFN and other resources, Karp is an important tool for accessing information. Searching on any expression, word form or lemma results in a display of every occurrence in all SweFN++ resources, except instances in the corpus. This gives, for example, an overview of different senses of polysemous words, in which resources they have been entered and how. Thus, we can see which SweFN frames are evoked by different senses of a word, we can see synonymous words in Swesaurus (Borin and Forsberg, 2014), the morphology of the word as well as multiword units containing this word in SALDO, samples of sentences from Korp where the chosen word occurs, and constructions in Swedish Constructicon which use this word (Lyngfelt et al., 2012).

SweFN developers use Karp to find SALDO entries that evoke a particular frame, SweCcn developers use Karp to find frames evoked by constructions, or constructions that evoke frames. Figure 2 shows an example of a view in Karp. In this particular view SweFN and SweCcn resources were selected, but other choices are also possible. The combination of searches shown here are in turn for a certain construction or frame (two first boxes), for constructions that match a certain frame (third box). This particular search is for constructions that match `Similarity`, which here resulted in 14 different constructions, each of which contained potential patterns which in turn could be used to perform new searches in Korp. Finally, in the fourth box the search is for a particular SALDO sense, and in the fifth box for a certain LU. Searches for other types of units such as frame elements, etc. are also possible.
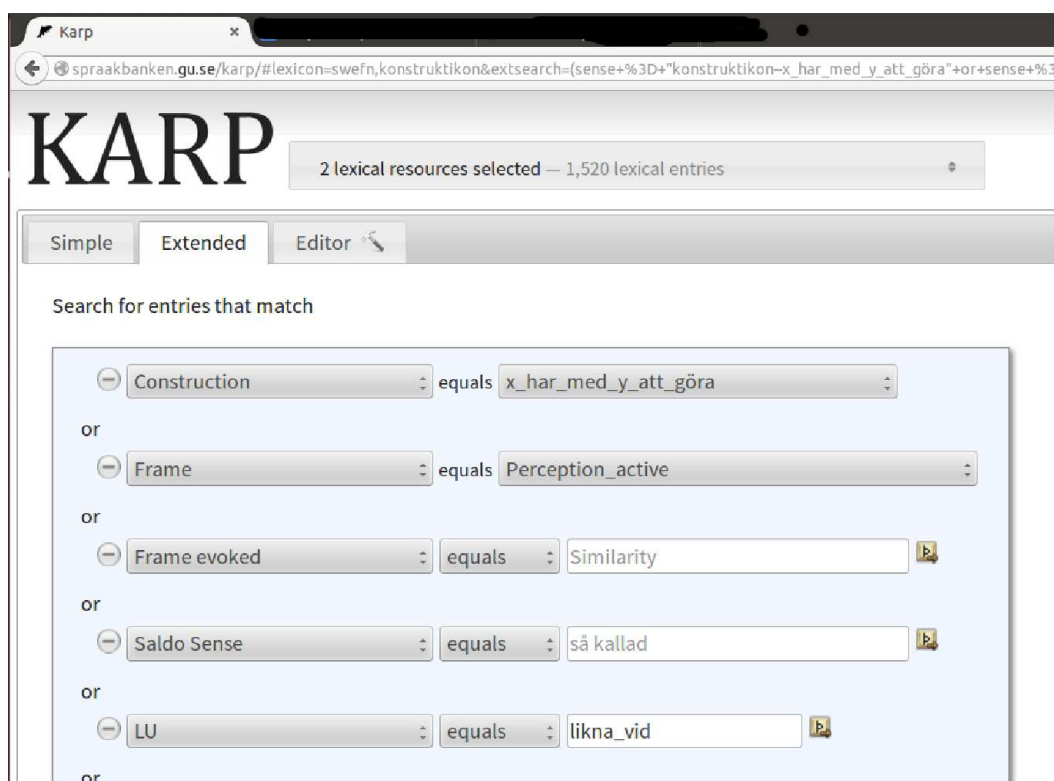

Figure 2: The Karp editing tool provides various functionalities to extract information from a number of different lexical resources. The combination of searches above is selected to illustrate the variety of possibilities in Karp.

4.4 Korp

The Korp corpora and search interface serve several purposes in the creation of SweFN. The coverage of lexical variation found in corpora is much larger than the variation we find in a lexicon and this helps in defining senses of polysemous words. From the corpora, example sentences are extracted to illustrate valence structures of LUs evoking frames. Korp extended search allows searches that combine SweFN LUs and syntactic structures of SweCcn constructions. The Related Words function provides a method of easily expanding the set of LUs populating a frame and giving easy access to example sentences where lexical variations are observed. Word Picture offers guidance in disambiguation as of LUs well as in analyzing semantic and syntactic structures.

Korp is a useful tool to check for compatibility with Swedish language and culture. Extended searches help us modify BFN frames and create new frames. There are two situations when BFN frames have been modified for SweFN (Heppin and Gronostaj, 2014): (1) the BFN frames are not suitable because of linguistic or cultural differences. For example the BFN frame `Jury_deliberation` has been redefined to `Deliberation` in SweFN. In `Deliberation` the FE corresponding to the FE JURY in BFN is changed to DELIBERATION_GROUP seeing that there is no jury in the Swedish legal process and a more general frame is appropriate as it covers deliberations in different kinds of legal systems; (2) the BFN frames are too general for our purposes, for example `Sound_makers` in BFN corresponds to two more specific frames in SweFN: `Noise_makers` and `Musical_instruments`. Completely new frames have also been created when there is a need for a frame not yet created for BFN. SweFN, for example, has a greater emphasis on nominal LUs than framenets for other languages. Therefore, frames such as `Animals`, `Countries`, and `Plants` have been created.

After determining the appropriate pairing of SALDO units and SweFN frames, searches are made for example sentences manifesting these LUs in the Korp corpora. The sentences we aim to find should have a variation of valence structure to give a broad overall picture of the LU patterns.

bygger (verb)

Subjekt		bygger	Objekt			Adverbial	
1. film	699		1. bo	459		1. nu	491
2. system	194		2. hus	270		2. ofta	227
3. undersökning	145		3. bro	120		3. just nu	117
4. metod	130		4. mur	79		4. mycket	86
5. verksamhet	154		5. fabrik	80		5. under tvåveckorsperioden	30
6. rapport	131		6. bostad	90		6. i tur^2	54
7. beräkning	86		7. bil	118		7. i tur	54
8. manus	80		8. larvbon	34		8. på roman	19
9. artikel	100		9. hyresrätt	46		9. på svar	18
10. resonemang	74		10. koja	33		10. på antagande	14
11. libretto	46		11. båt	65		11. på bok	15
12. bok	147		12. lägenhet	70		12. samtidig	97
13. studie	120		13. nätverk	52		13. i ställ	46
14. bok^2	130		14. muskel	46		14. i ställe	46
15. modell	97		15. tunnel	39		15. på enkät	10

Figure 3: Word picture from Korp of the verb bygga 'build' in present tense, e.g. *bygger*. The columns display from left to right subjects, objects, and adverbials. The number to the right in each column is the frequency of the collocation in Korp.

Word Picture is useful when taking a starting point in individual, polysemous words, to determine which frames are evoked by the different senses. In figure 3 items, which are listed in subject and object

positions respectively, highlight two different senses of the verb *bygga* 'build', one abstract and one concrete sense. The nouns found in subject position, such as *film* 'film', *system* 'system', *undersökning* 'examination', *metod* 'method', *rapport* 'report', etc., occur with the sense of *bygga* 'build' which is typically found in an abstract intransitive construction with the preposition *på* 'on' as in 'founded on', 'built on', or 'based on'. This sense evokes the `Use_as_a_starting_point` frame. The nouns in the object position, such as *hus* 'house' and *bro* 'bridge', collocate with the agentive verb *bygga* 'build' in the concrete sense of 'construct' or 'erect', which evokes the `Building` frame (Heppin and Gronostaj, 2014).

5   Consistency checks and automatic extension of the data

There is no gold standard to evaluate the quality of SweFN against as there is no other comparable resource. FrameNet-like resources for other languages are constructed with different foci and under different conditions. However, there is a constant assessment of the correctness of the resources built into the workflow and ongoing consistency checks to avoid inconsistency between resources. The Karp tool gives error messages, for example when SALDO entries are listed in more than one frame. Other types of checks are run with certain intervals, for example to see if there are annotation tags which do not follow the standard format. Confronted with different types of error messages the developers go back to the frames in question to revise the contents of the frame, such as which LUs are said to evoke the frame, or the choice of and annotation of example sentences.

One part of the work is directed towards developing computational methods to facilitate the manual construction of SweFN. We have so far focused on three tasks: (1) semantic role labeling (SRL) (Johansson et al., 2012); (2) automatic sentence extraction, i.e. finding example sentences with varied syntactic and semantic complexities (Pilán et al., 2013); (3) automatic expansion of the SweFN lexicon to determine which frame is evoked by a given word by combining statistical and rule-based methods based on SALDO descriptors and extracted information from Korp (Johansson, 2014).

6   Conclusions

The building of one big macro-resource for Swedish language technology, where the individual resources interact with and enhance each other, provides a unique overview of the Swedish language. One search on a lexical expression results in a list of descriptions from all of the separate resources. The information derived is not only useful for the end user, but also for the continuing work on all parts of the linguistic macro-structure.

We have here focused on how two language technology resources, SALDO and SweCcn, are exploited in the development of SweFN, but also on how these resources enhance each other and other resources. We mainly address the manual perspectives of the workflow, illustrating what data may derive from the different resources, how this data may be used to facilitate work, and how the contents of one resource may reappear in the contents of another. We have given a sketch of the language technology tools with the aim to reveal their potential importance in the development of SweFN.

The construction of SweFN, and even more so the construction of a macro-resource such as SweFN++, will continue to develop in the foreseeable future. New insights as well as new problems will continue to give rise to changes.

Acknowledgements

References

Malin Ahlberg, Lars Borin, Markus Forsberg, Martin Hammarstedt, Leif-Jöran Olsson, Olof Olsson, Johan Roxendal, and Jonatan Uppström. 2013. Korp and Karp a bestiary of language resources: the research infrastructure of Språkbanken. In *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2013), Oslo University, Norway. NEALT Proceedings Series 16*, number 16, pages 429–433.

Héctor Martínez Alonso, Bolette Sandford Pedersen, and Núria Bel. 2013. Annotation of regular polysemy and underspecification. In *ACL (2)*, pages 725–730. The Association for Computer Linguistics.

Linnéa Bäckström, Benjamin Lyngfelt, and Emma Sköldberg. 2014. Towards interlingual constructicography. on correspondence between constructicon resources for English and Swedish. *Constructions and Frames*, 6(1):9–32. John Benjamins Publishing Company.

Anderson Bertoldi and Rove Luiza de Oliveira Chishman. 2011. The limits of using FrameNet frames to build a legal ontology. In *CEUR Workshop Proceedings*, volume 776, pages 207–212.

Lars Borin and Markus Forsberg. 2014. Swesaurus; or, The Frankenstein approach to Wordnet construction. In *Proceedings of the Seventh Global WordNet Conference (GWC 2014)*.

Lars Borin, Dana Dannélls, Markus Forsberg, Maria Toporowska Gronostaj, and Dimitrios Kokkinakis. 2010. The past meets the present in Swedish FrameNet++. In *Proceedings of the 14th EURALEX International Congress*, pages 269–281.

Lars Borin, Markus Forsberg, and Johan Roxendal. 2012. Korp the corpus infrastructure of Språkbanken. In *Proceedings of LREC 2012*. Istanbul: ELRA.

Lars Borin, Markus Forsberg, and Lennart Lönngren. 2013a. SALDO: a touch of yin to WordNet's yang. *Language Resources and Evaluation*, 47(4):1191–1211.

Lars Borin, Markus Forsberg, Leif-Jöran Olsson, Olof Olsson, and Jonatan Uppström. 2013b. The lexical editing system of Karp. In *Proceedings of the eLex 2013 conference*, pages 503–516, Tallin.

Aljoscha Burchardt, Katrin Erk, Anette Frank, Andrea Kowalski, Sebastian Padó, and Manfred Pinkal, 2009. *Multilingual FrameNets in computational lexicography*, chapter Using FrameNet for the semantic analysis of German annotation, representation, and annotation. Berlin: Mouton de Gryter.

Dipanjan Das, Desai Chen, André F. T. Martins, Nathan Schneider, and Noah A. Smith. 2014. Frame semantic parsing. *Computational Linguistics*, 40(1):9–56.

Anna Ehrlemark. 2014. Ramar och konstruktioner – en kärlekshistoria [Frames and constructions – a love story]. Department of Swedish, University of Gothenburg. GU-ISS 2014-01.

Charles J. Fillmore, Miriam R.L. Petruck, Josef Ruppenhofer, and Abby Wright. 2003. FrameNet in Action: The Case of Attaching. *IJL*, 16(3):297–332, September.

Charles J. Fillmore, Russell Lee-Goldman, and Russell Rhomieux, 2012. *Sign-based construction grammar*, chapter The FrameNet constructicon. Stanford: CSLI.

Charles J. Fillmore. 1985. Frames and the semantics of understanding. *Quaderni di Semantica*, 6(2):222–254.

Gil Francopoulo, Monte George, Nicoletta Calzolari, Monica Monachini, Nuria Bel, Mandy Pet, and Claudia Soria. 2006. LMF for Multilingual, Specialized Lexicons. In *Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC)*, pages 233–236.

Karin Friberg Heppin and Maria Toporowska Gronostaj. 2014. Exploiting FrameNet for Swedish: Mismatch? *Constructions and Frames*, 6(1):51–71. John Benjamins Publishing Company.

Richard Johansson, Karin Friberg Heppin, and Dimitrios Kokkinakis. 2012. Semantic role labeling with the Swedish FrameNet. In *Proceedings of the 8th Conference on International Language Resources and Evaluation (LREC)*, pages 3697–3700, Istanbul, Turkey.

Richard Johansson. 2014. Automatic expansion of the Swedish FrameNet lexicon – Comparing and combining lexicon-based and corpus-based methods. *Constructions and Frames*, 6(1):91–112. John Benjamins Publishing Company.

Ora Lassila and Ralph Swick. 1999. Resource Description Framework (RDF). Model and Syntax Specification. Technical report, W3C. `http://www.w3.org/TR/REC-rdf-syntax`.

Benjamin Lyngfelt, Lars Borin, Markus Forsberg, Julia Prentice, Rudolf Rydstedt, Emma Sköldberg, and Sofia Tingsell. 2012. Adding a Constructicon to the Swedish resource network of Språkbanken. In *Proceedings of KONVENS 2012*, Vienna. LexSem workshop.

Alessandro Moschitti, Paul Morarescu, and Sanda M. Harabagiu. 2003. Open domain information extraction via automatic semantic labeling. In *Proceedings of the 16th International FLAIRS Conference*, pages 397–401.

Joakim Nivre, Johan Hall, Jens Nilsson, Atanas Chanev, Glsen Eryigit, Sandra Kbler, Svetoslav Marinov, and Erwin Marsi. 2007. MaltParser: A language-independent system for data-driven dependency parsing. *Natural Language Engineering*, 13(2):95–135.

Kyoko Hirose Ohara. 2013. Toward constructicon building for Japanese in Japanese FrameNet. *Veredas: Frame Semantics and Its Technological Applications*, 17(1):11–28.

Ildikó Pilán, Elena Volodina, and Richard Johansson. 2013. Automatic selection of suitable sentences for language learning exercises. In *20 Years of EUROCALL: Learning from the Past, Looking to the Future. 2013 EUROCALL Conference, 11th to 14th September 2013 Evora, Portugal, Proceedings.*, pages 218–225.

Michael Roth and Anette Frank. 2009. A NLG-based application for walking directions. In *Proceedings of the 47th ACL and the 4th IJCNLP Conference (Software Demonstrations)*, pages 37–40.

Emma Sköldberg, Linnéa Bäckström, Lars Borin, Markus Forsberg, Benjamin Lyngfelt, Leif-Jöran Olsson, Julia Prentice, Rudolf Rydstedt, Sofia Tingsell, and Jonatan Uppström. 2013. Between grammars and dictionaries: a Swedish Constructicon. In *Proceedings of the eLex 2013 conference*, pages 310–327, Tallin.

Carlos Subirats, 2009. *Multilingual FrameNets in Computational Lexicography*, chapter Spanish FrameNet: a frame-semantic analysis of the Spanish lexicon. Berlin: Mouton de Gryter.

Tiago Timponi Torrent, Maria Margarida Martins Salomão, Ely Edison da Silva Matos, Maucha Andrade Ganomal, Júlia Gonçalves, Bruno Pereira de Souza, Daniela Simões, and Simone Rodrigues Peron-Corrêa. 2014. Multilingual lexicographic annotation for domain-specific electronic dictionaries: the Copa 2014 FrameNet Brasil project. *Constructions and Frames*, 6(1):72–90. John Benjamins Publishing Company.

Tiago Timpioni Torrent. 2013. Behind the labels: Criteria for defining analytical categories in FrameNet Brasil. *Veredas: Frame Semantics and Its Technological Applications*, 17(1):44–65.