

# Guests and hosts - Achieving clarity in vagueness in Swedish FrameNet

Karin Friberg Heppin

Språkbanken, Department of Swedish

University of Gothenburg

karin.friberg.heppin@svenska.gu.se

## 1. Introduction

In the FrameNet formalism lexical units (LUs) are pairings of words and semantic frames, each LU representing one sense. If a word evokes more than one frame it should be represented by several LUs, one for each sense. Often there is a straightforward relation between frames and possible senses of words; however, cases of vagueness or regular polysemy need more elaboration (Friberg Heppin & Dannélls 2015). During the construction of Swedish FrameNet (SweFN)<sup>1</sup> polysemous and vague words have forced a systematic approach to the lumping or splitting of lexical entries and semantic frames. We suggest creating explicit links for LUs appearing as Guest LUs in host frames, other than their primarily evoked frames. This would be a method of creating clarity in vagueness and in regular polysemy without resorting to undue splitting of lexical entries in the lexicon.

## 2. Frames evoked

An assumption made in the early development stages of SweFN was that lexical entries from the reference lexicon SALDO (Borin et al., 2013) would be used as LUs. Each SALDO entry would be allowed to evoke one semantic frame only. If an entry, none the less, appeared to evoke more than one frame it entailed an additional sense and therefore required a request for a new entry in SALDO. This is in contrast to the original English language FrameNet (FN) which is not linked to any existing lexicon and therefore does not have the same restrictions in how many frames a word may evoke.

Ruppenhofer et al. (2010) describe the Non-perspectivalized frames which exist in FN. A frame of this type differs from other frames in that they contain a diversity of LUs sharing a certain scene as background, but which do not necessarily have consistent semantic types. An example is the `Education_teaching`<sup>2</sup> frame, which is evoked by LUs such as *study.v*, *teach.v*, *training.n*, *educational.a*, and *teacher.n*. The purpose of Non-perspectivalized frames in FN is described as being a time-saving measure (Ruppenhofer et al., 2010).

As Non-perspectivalized frames are evoked by LUs according to background, rather than semantic type, there is sometimes a clash with other frames evoked by more uniform sets of LUs. Such is the case in FN for the noun *teacher* which evokes the frame `Education_teaching` as well as `People_by_vocation`, a frame populated by nouns for vocations. According to the FrameNet formalism evoking two frames entails two senses of the word, however, without explicit links to a lexicon it is not a problem. The corresponding situation for SweFN is more problematic as it demands two entries in SALDO for the equivalent Swedish word *lärare*, a split which is difficult to justify from other than a FrameNet

perspective.

While Non-perspectivalized frames are evoked by LUs not always consistent in the semantic roles they carry along, Guest LUs are LUs used in a manner which evokes frames other than their primary ones. These frames are called host frames. The interpretation of a Guest LU depends on both primary and host frames. Not allowing or considering Guest LUs in the construction of a FrameNet resource, while utilizing the entries of an existing lexicon, also tends to give bias to the splitting, rather than the lumping, point of view regarding lexical entries, as we saw in the case of *lärare* above. Work on constructing a particular frame includes the phase of populating it with LUs. Encountering an entry in a lexicon, or a word in a corpus sentence, it is tempting to list it as an LU in any frame it could be said to evoke. However, the potential of an entity to evoke a frame does not mean that this is the only possible use nor the primary use.

Hanks (2013) describes words as having meaning potentials in that different senses are activated in different contexts, something which does not entail that the word in question has several distinctive senses. The fuzziness is a strength rather than a flaw, as it makes language flexible. It is not always desirable to be specific. As stated by Wierzbicka (1984) the aim must sometimes be vagueness:

An adequate definition of a vague concept must aim at precision in vagueness – it must aim at PRECISELY that level of vagueness which characterizes the concept itself. (Wierzbicka, 1984):210

## 3. Related meaning potentials

There are groups of words which may take on closely related meanings. Take the example of words describing the relationship between persons and geographical places such as Canadian and Londoner describing persons with certain geographical origin. However, the same word may also describe where a person resides or where he or she is a citizen, which is not necessarily the same place. When stating a person's nationality or city it may be an advantage to be vague in this respect. In FN and SweFN there are three frames which may be evoked by words for origin/residence/citizenship: `People_by_origin`, `Residence`, and `People_by_jurisdiction`, all inheriting directly from the `People` frame.

Most words denoting people in relation to geographic areas could evoke all three frames. However, a few evoke only one e.g.: *malmöbo* (Malmö-liver) 'Malmö resident' evoking `Residence`, and *svenskfödd* (Swedish-born) 'born in Sweden' evoking `People by origin`. For the words which potentially evoke all three frames it is often desirable to maintain vagueness, leaving it to context to determine which meaning potentials should be realized. This may be solved by creating a new a frame on an intermediate level, inheriting from `People` and itself having its LUs hosted by the other three frames. The parent

<sup>1</sup> <<https://spraakbanken.gu.se/eng/swefn/>>

<sup>2</sup> In this paper frame names are written in Courier.

frame in this case could have a name such as *People\_by\_locale*. LUs such as *malmöbo* and *svenskfödd* should not populate this frame, but rather only the frames that they do evoke, meaning they should be moved to a lower level in the frame hierarchy. An explicit guest-host relationship of this type could be seen as an inverse inheritance. Instead of LUs of the child frame having the potential of evoking the parent frame, the child frame would host LUs of the parent frame.

#### 4. Diverse meaning potentials

A group of words which is often used underspecified, having several meaning potentials of diverse semantic types, are words denoting institutions/businesses/organizations, including the activities and people within. For some of these we have Non-perspectivalized frames such as for the LU *skola* ‘school’ evoking *Education teaching*. Others, like *bank* ‘bank’, do not.

We can take a look at how the noun *skola* is presented in *Svensk ordbok*, a monolingual Swedish dictionary published by the Swedish Academy (Allén et al., 2009). This noun is described by one main sense and four sub-senses. Below we have paired these senses with the frames they potentially evoke followed by part of the frame description:

- Institution where education is performed - evokes *Institutions* “permanent organizations [...] with a public character”

1. with focus on the activities performed within the institution – evokes *Education teaching* “words referring to teaching and the participants in teaching”

2. with focus on the building where the education is performed – evokes *Buildings* “contains words which name permanent fixed structures forming an enclosure and providing protection from the elements”

3. with focus on the collective of persons working with/attending activities within a certain institution – evokes *Aggregate* “nouns denoting *Aggregates of Individuals*”

#### Acknowledgements

The work presented was made possible by the Swedish Research Council (grant agreement 2010-6013), the Bank of Sweden Tercentenary Foundation (grant agreement P120076:1), and by the University of Gothenburg through its support of the Centre for Language Technology and of Språkbanken.

#### Reference

S. Allén, D. Berg, S. Berg, M. Gellerstam, L. Holmer, A-K. Hult, S. Lindstrand, S. Lövfors, S-G. Malmgren, C. Sjögren, E. Sköldberg, L. Tegner, and M. Toporowska Gronostaj, editors. 2009. *Svensk ordbok utgiven av Svenska Akademien*. 1-2.

L. Borin, M. Forsberg, and L. Lönngrén. 2013. SALDO: a touch of yin to WordNet’s yang. *Language Resources and Evaluation*, 47(4):1191–1211.

K. Friberg Heppin and D. Dannélls. 2015. Polysemy and

4. other organization teaching a particular skill or subject – evokes *Organization* “intentionally formed human social groups [...] with some definite structure and Members”

Words with different meaning potentials are often used underspecified, with more than one sense. This is seen in example (1), including the word *skola*. Here the visitor, *Jag* ‘I’, may be seen as visiting the persons, the activities, as well as the building of the school itself. Making separate entries in the lexicon for each potential, each becoming an LU evoking a different frame, would not catch the possibility of vagueness and the relations between the senses would be lost.

(1) Jag ska besöka en skola i Köpenhamn.  
‘I am going to visit a school in Copenhagen.’

For LUs evoking Non-perspectivalized frames, as well as LUs with potential meanings evoking one or more other frames, e.g. *bank*, which until now populates only the frame *Businesses* in SweFN, we suggest the same type of linking as for words with related meaning potentials: an explicit linking between all LUs or subsets of LUs of concerned frames and frames representing potential meanings or uses, allowing LUs in primary frames to be Guest LUs in one or more host frames simultaneously.

#### 5. Conclusion

The idea of Guest LUs and host frames is not a novel one, but it has not been utilized to its full potential. Explicit links in primary frames declaring all LUs or subsets of LUs within to be Guest LUs in host frames could be a powerful method for bringing clarity to possible interpretations of vague LUs. Not yet existing parent frames, such as *People\_by\_locale*, could be created in cases of related senses for this purpose.

questions of lumping or splitting in the construction of Swedish FrameNet. In *Proceedings of the Workshop on Semantic Resources and Semantic Annotation for Natural Language Processing and the Digital Humanities at NODALIDA*, pages 12-20. NEALT:

K. Friberg Heppin and M. Toporowska Gronostaj. 2014. Exploiting FrameNet for Swedish: Mismatch? *Constructions and Frames*, 6(1):52–72.

P. Hanks. 2013. *Lexical Analysis: Norms and Exploitations*. MIT Press, Cambridge, Mass.

J. Ruppenhofer, M. Ellsworth, M. Petruck, C. Johnson, and J. Scheffczyk. 2010. *FrameNet II: Extended Theory and Practice*. ICSI, Berkeley.

A. Wierzbicka. 1984. Cups and mugs: Lexicography and conceptual analysis. *Australian Journal of Linguistics*, 4(2):205–255.